FIG. 1

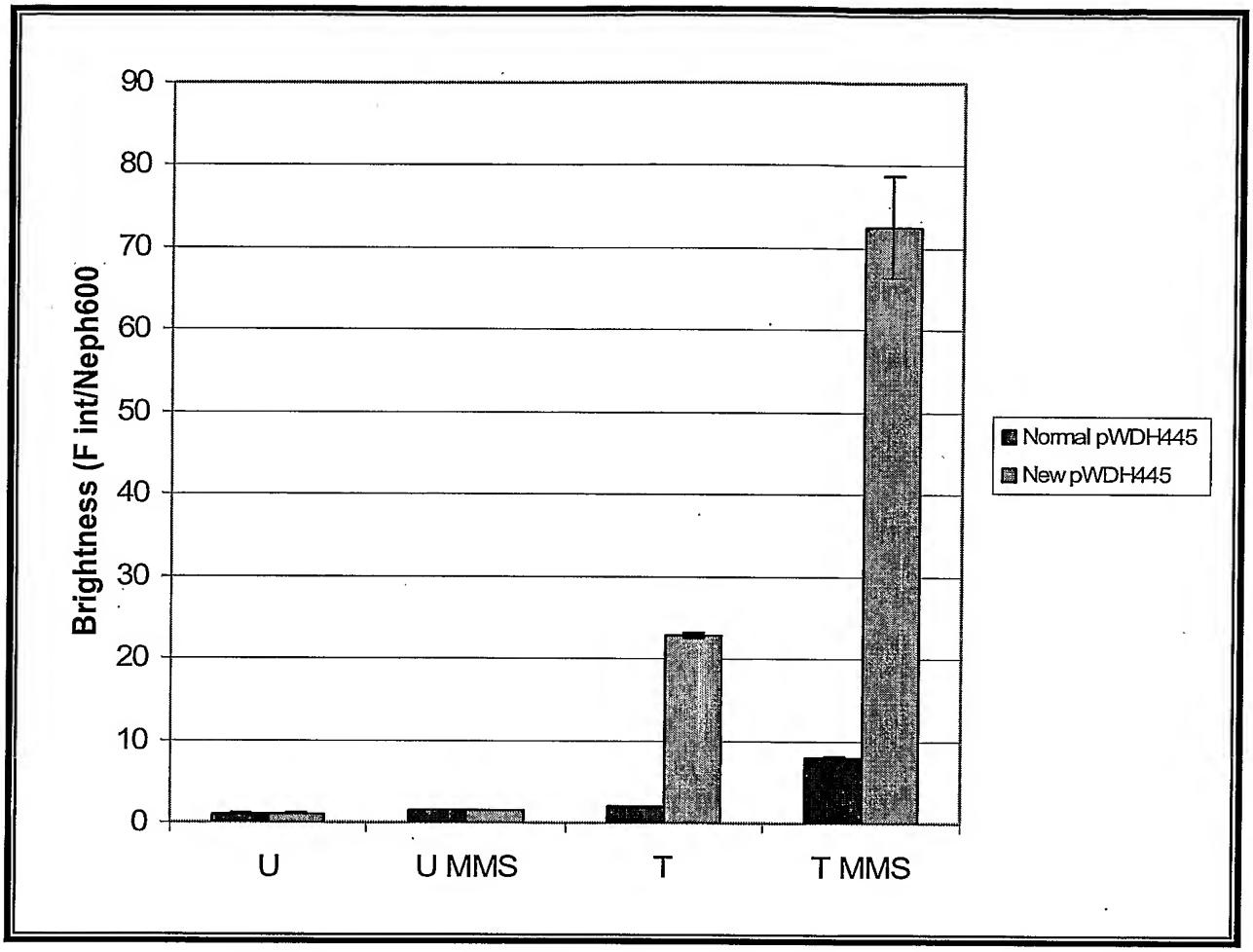

Brightness (F int/Neph600
90
80
70
60
50
40
30
20
10
0
U
U MMS
T
T MMS
Normal pWDH445
New pWDH445

FIG.2

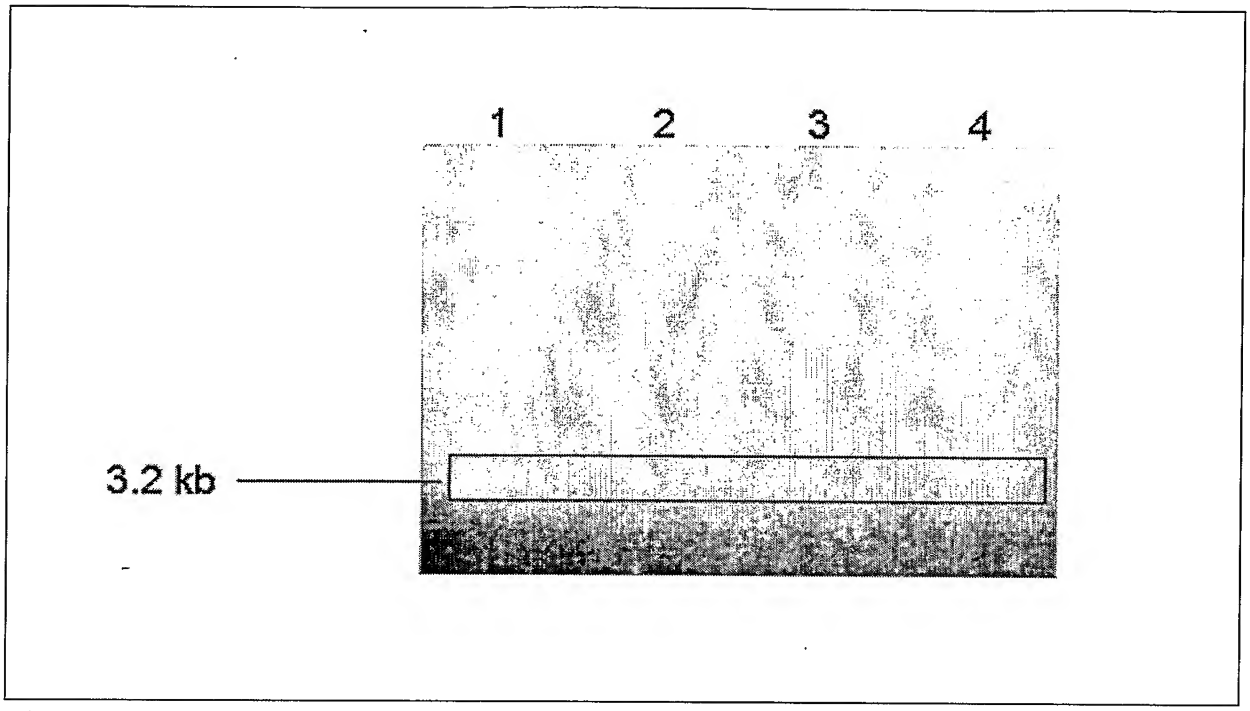


FIG. 3

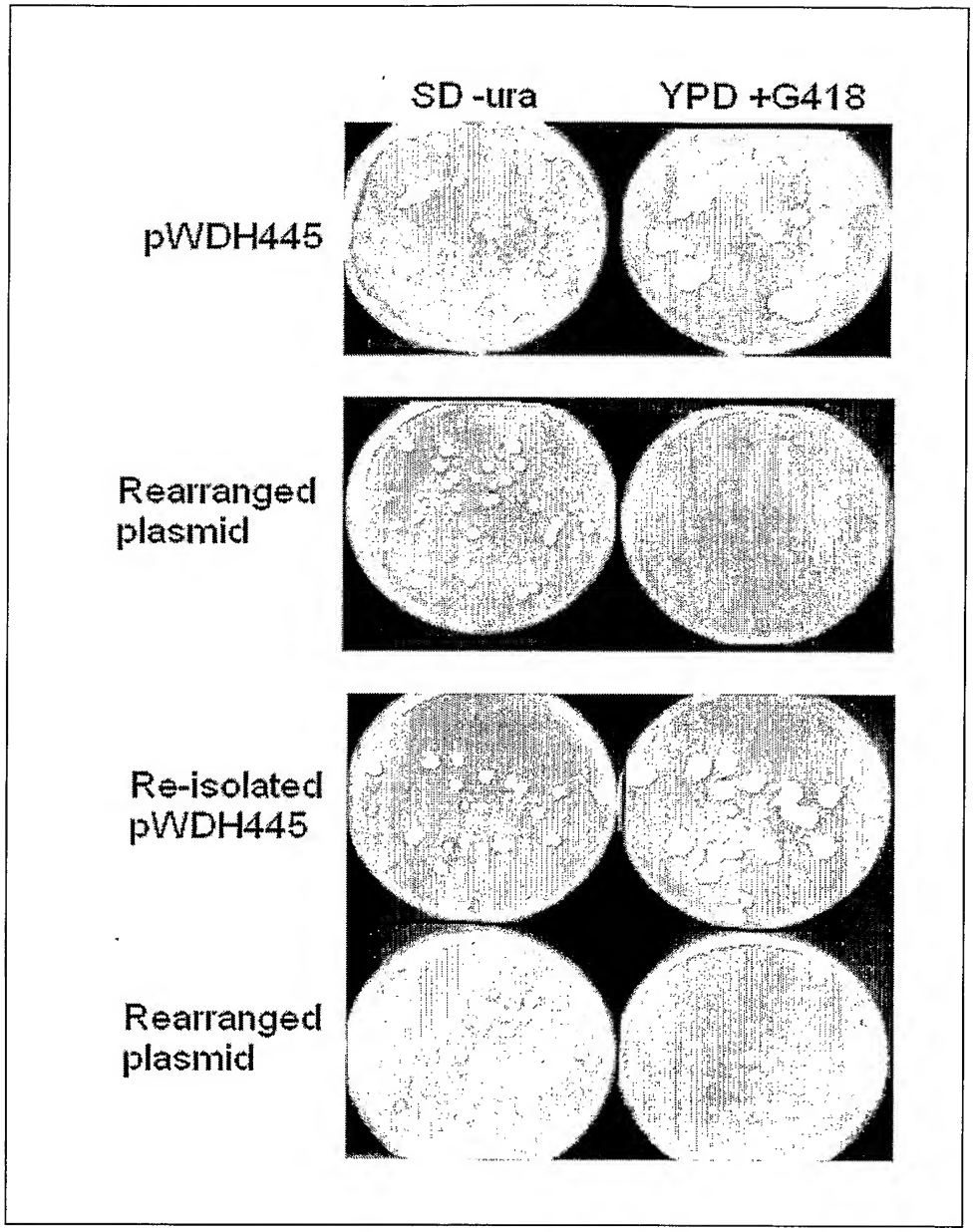

SD -ura
YPD +G418
pWDH445
Rearranged plasmid
Re-isolated pWDH445
Rearranged plasmid

FIG. 4

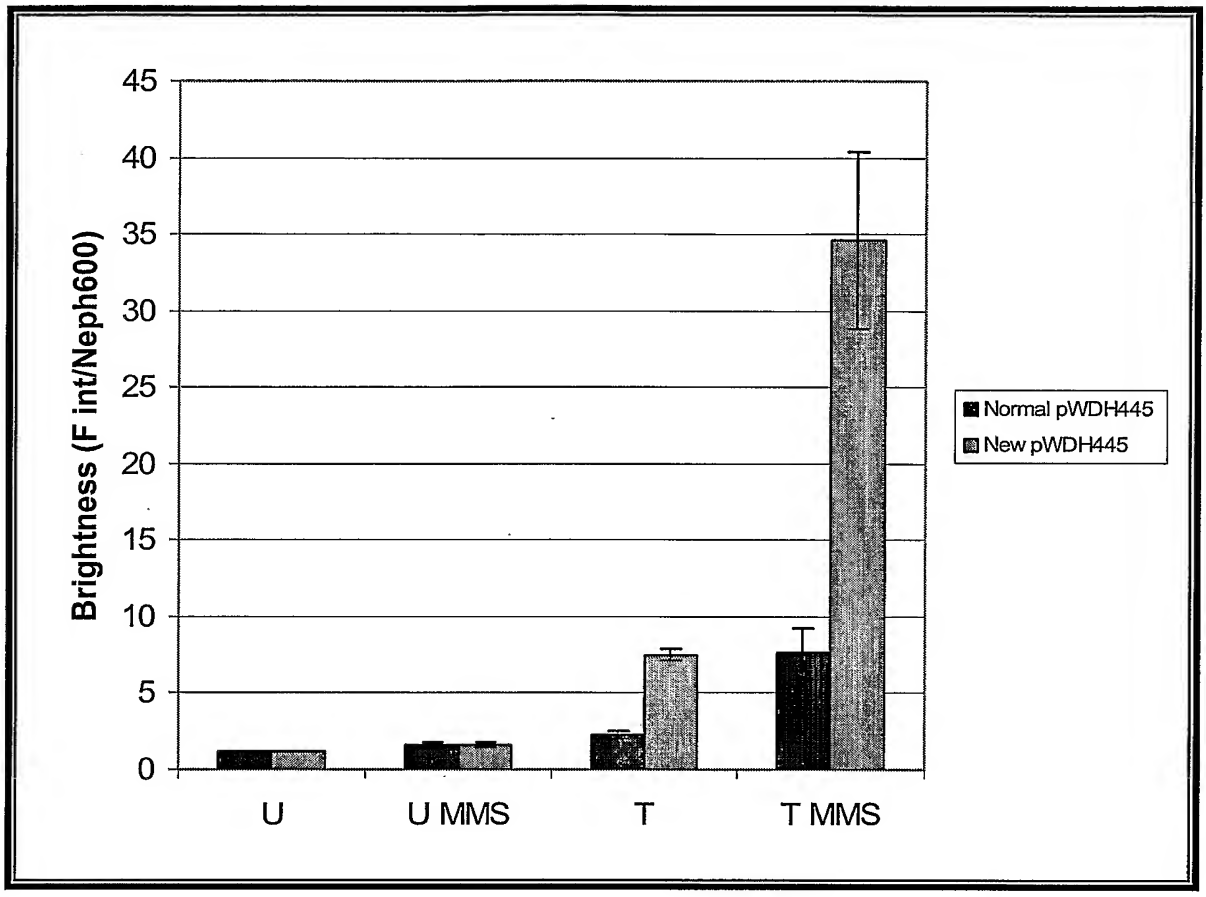

Brightness (F int/Neph600)
45
40
35
30
25
20
15
10
5
0
U
U MMS
T
T MMS
Normal pWDH445
New pWDH445

FIG. 5

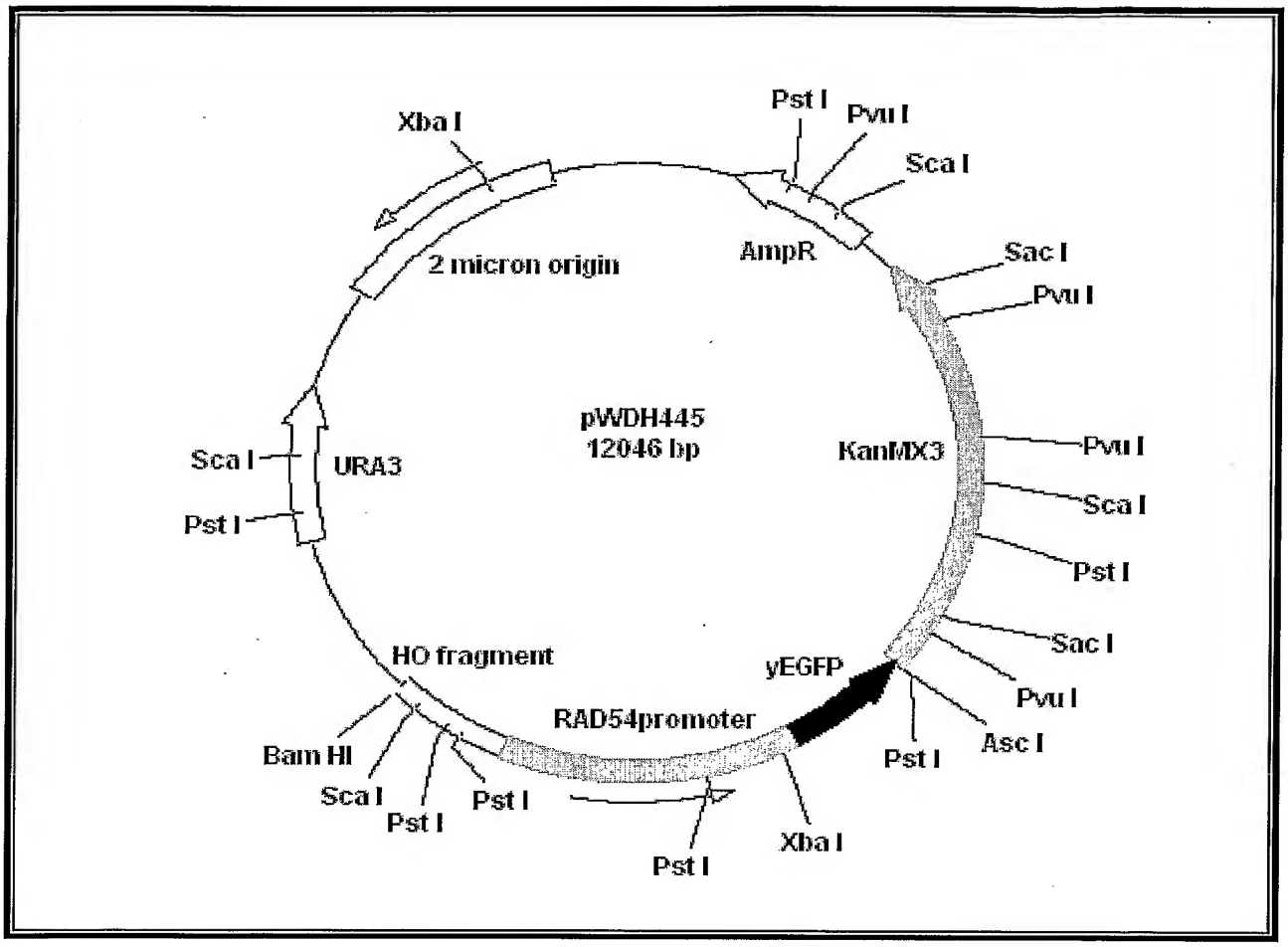

Xba I
Pst I
Pvu I
Sca I
AmpR
Sac I
Pvu I
2 micron origin
pWDH445
12046 bp
KanMX3
Pvu I
Sca I
Pst I
Sca I
URA3
Pst I
Sac I
Pvu I
HO fragment
yEGFP
RAD54promoter
Asc I
Pst I
Bam HI
Sca I
Pst I
Pst I
Pst I
Xba I

FIG. 6

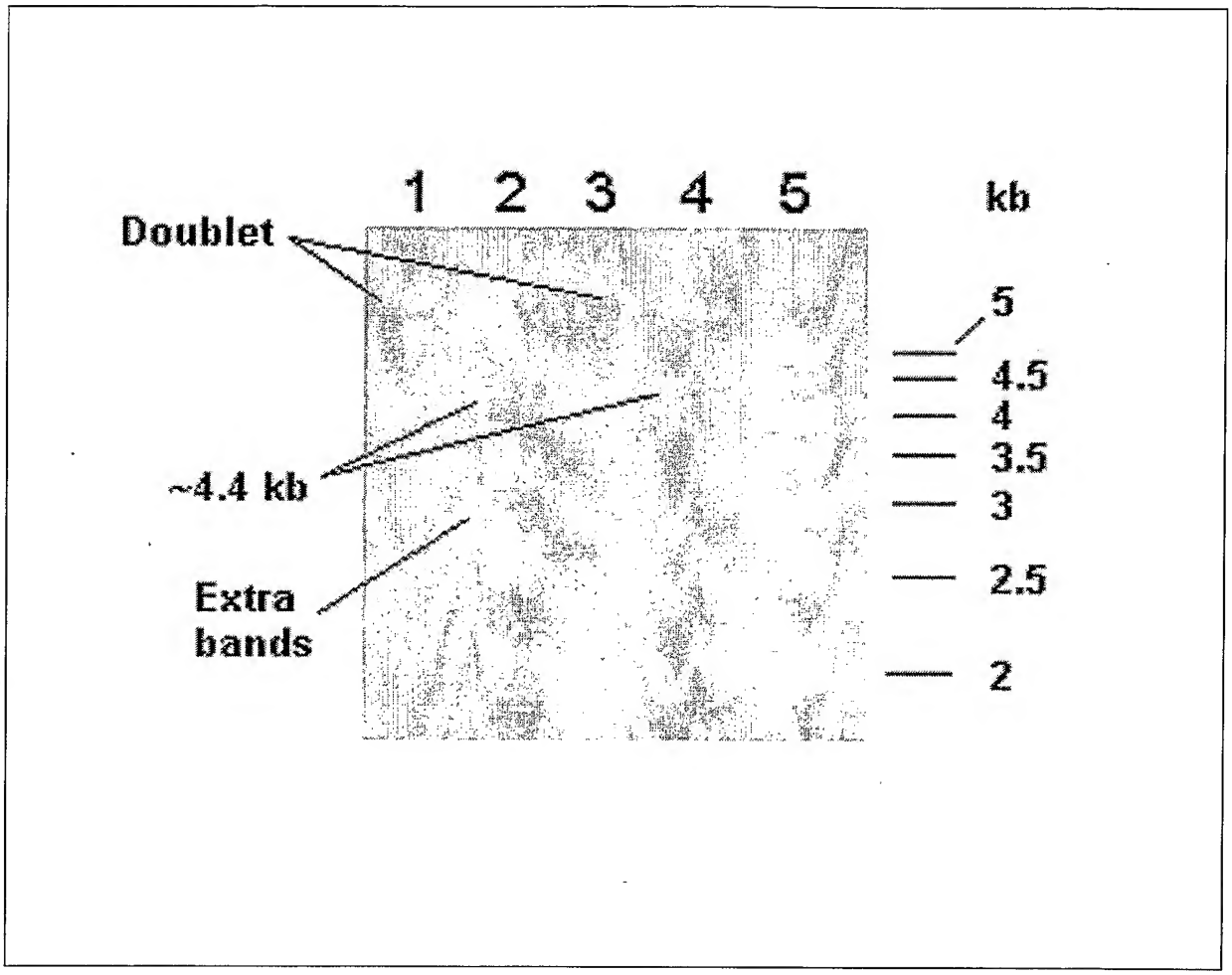


FIG. 7

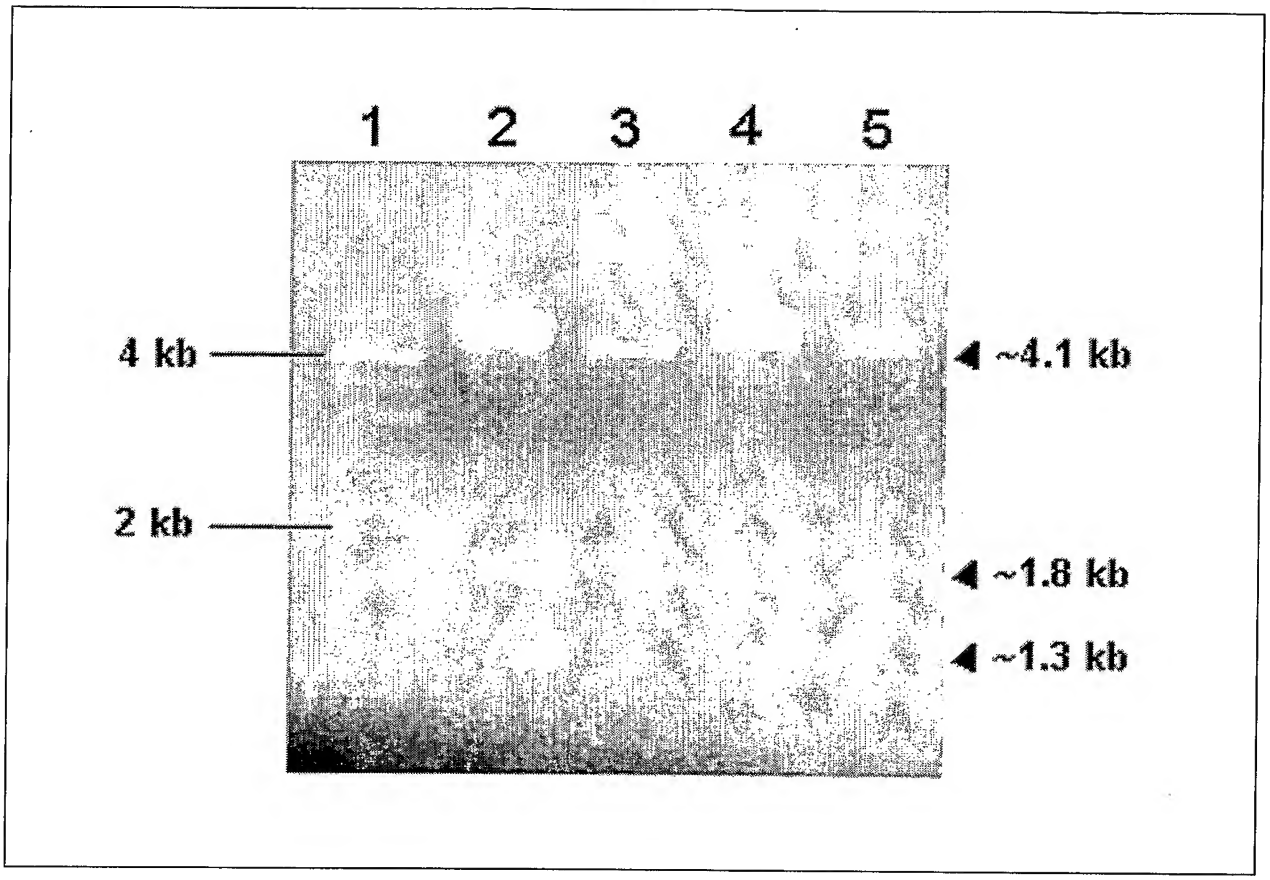


FIG. 8

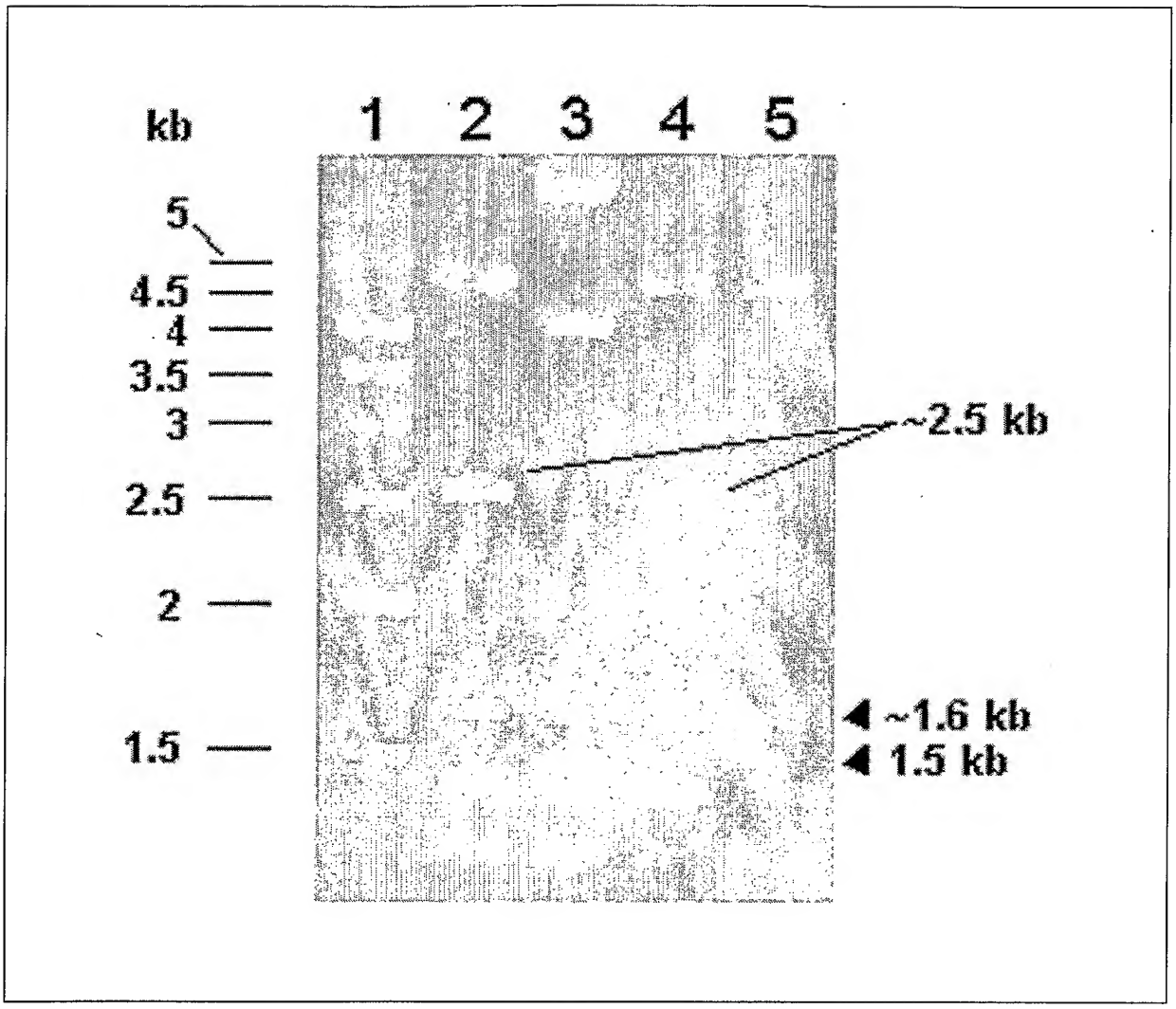


FIG.9

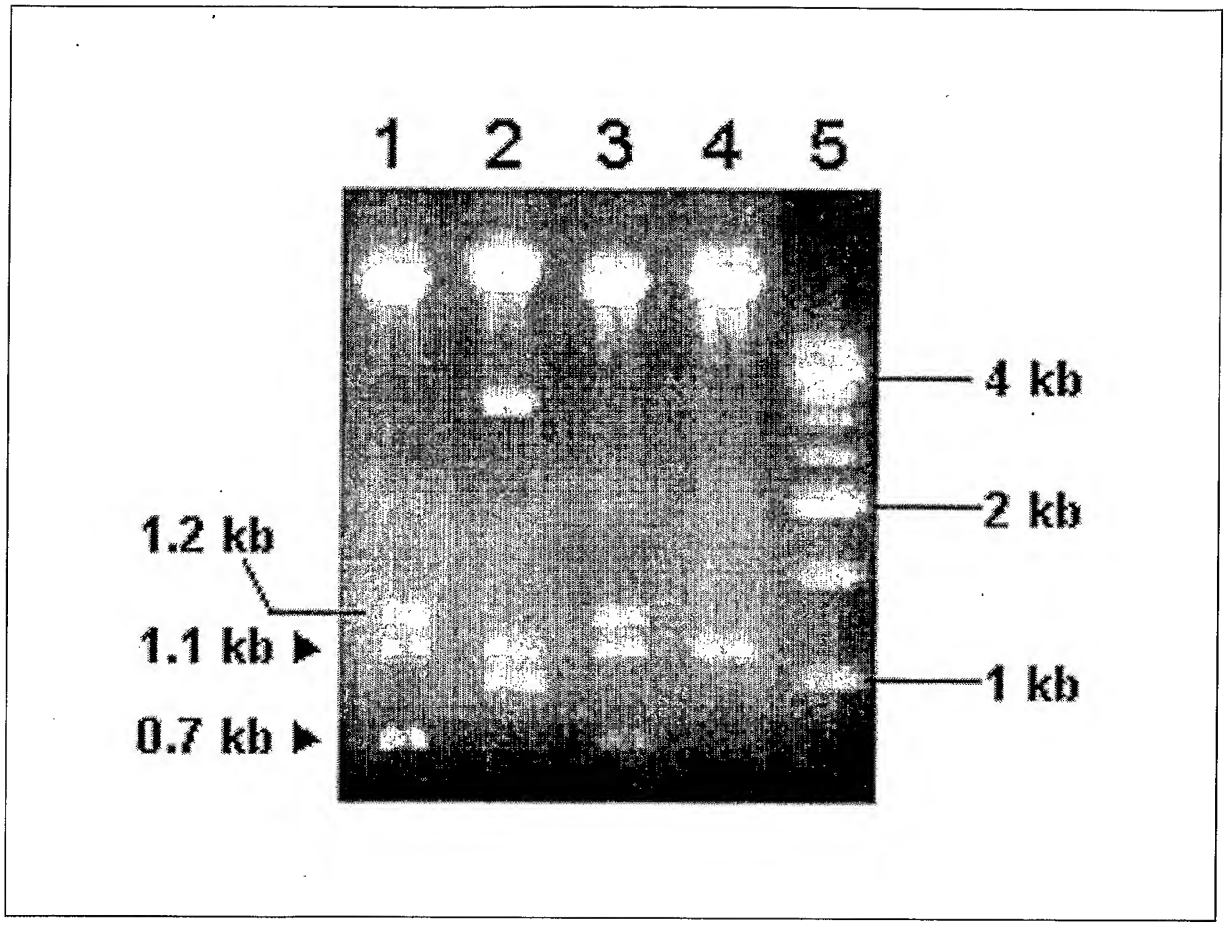


FIG. 10

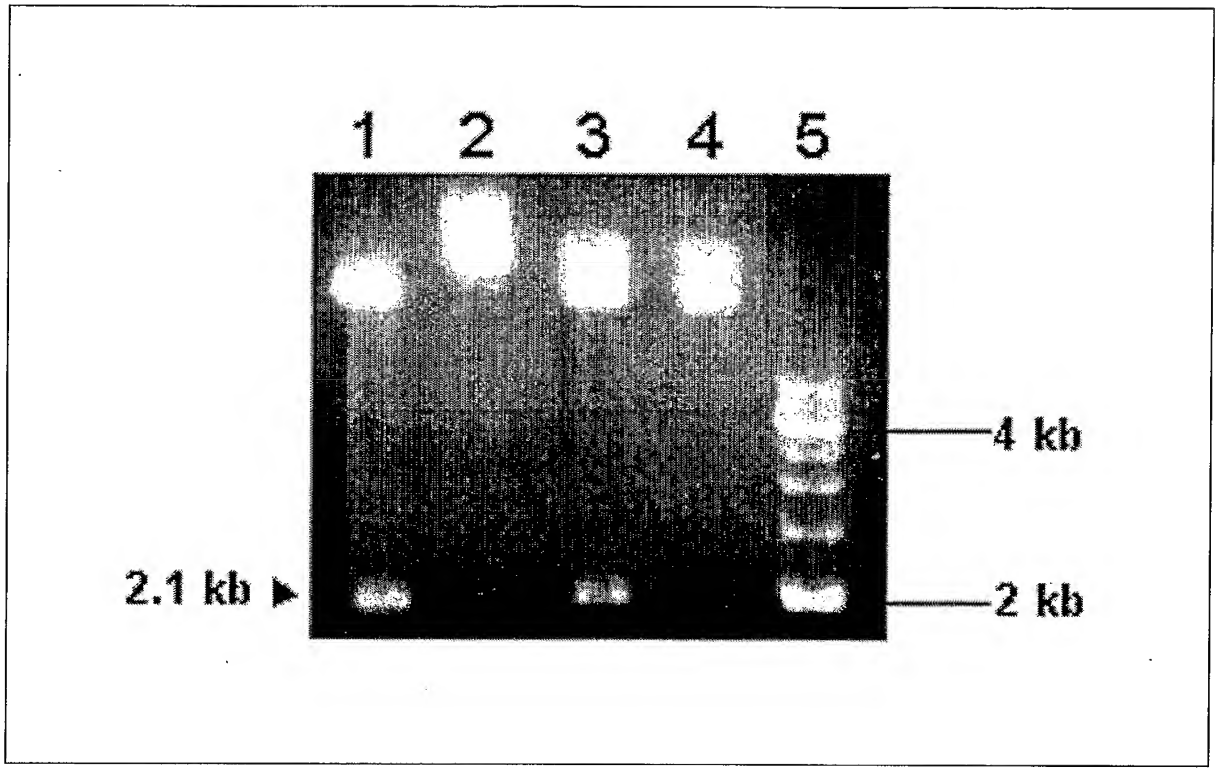


FIG. 11

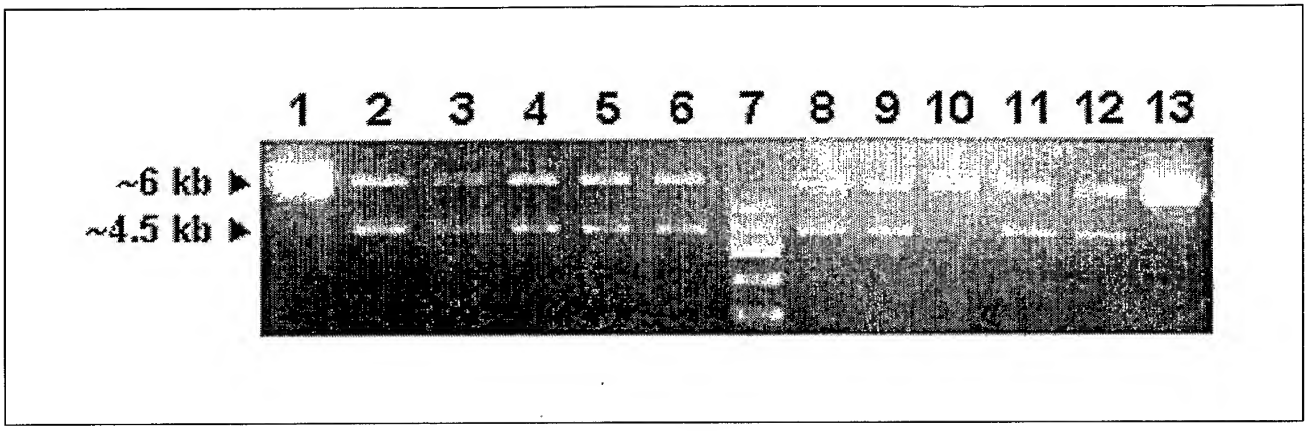


FIG. 12

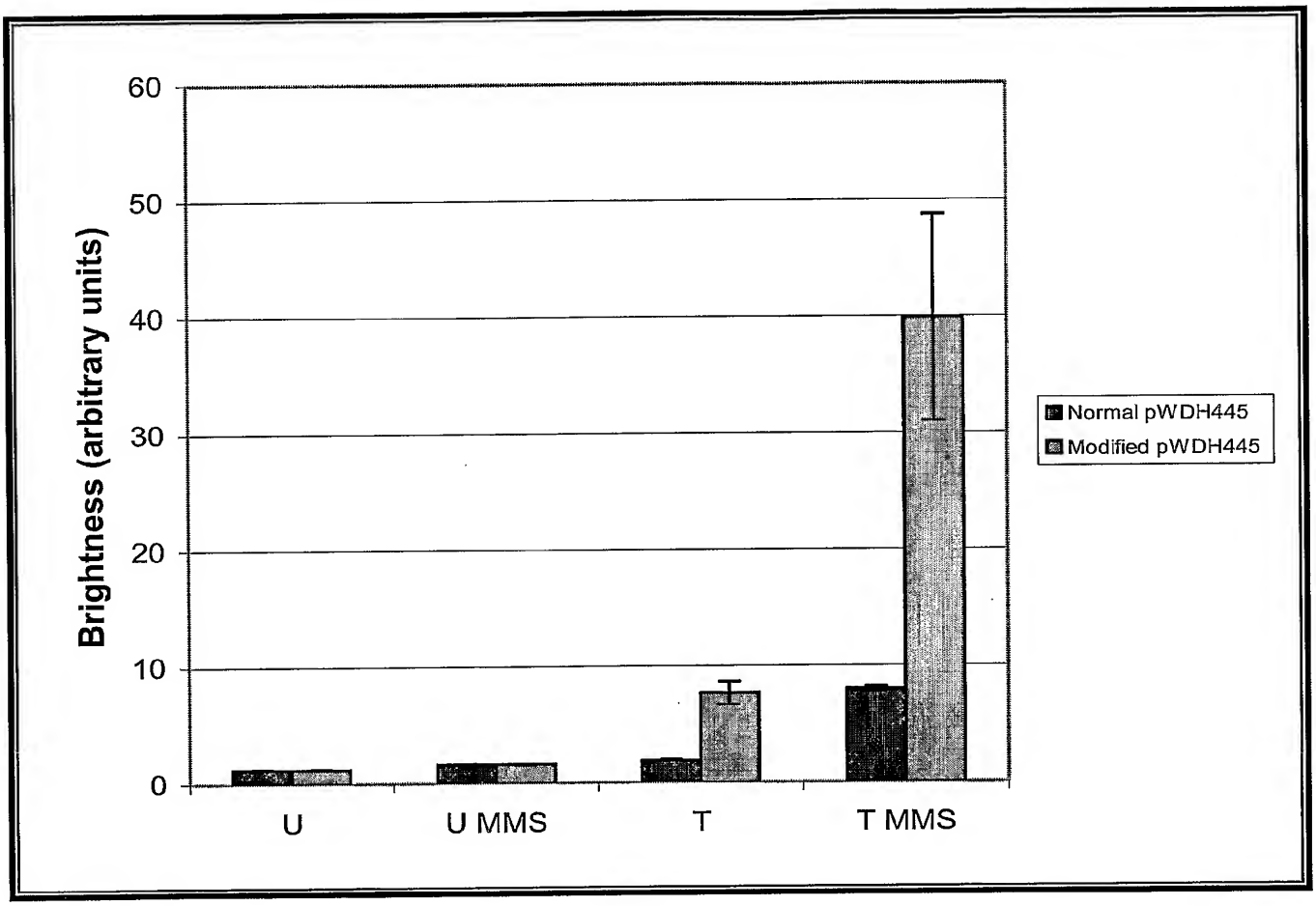

60
50
40
30
20
10
0
Brightness (arbitrary units)
U
U MMS
T
T MMS
Normal pWDH445
Modified pWDH445

FIG. 13

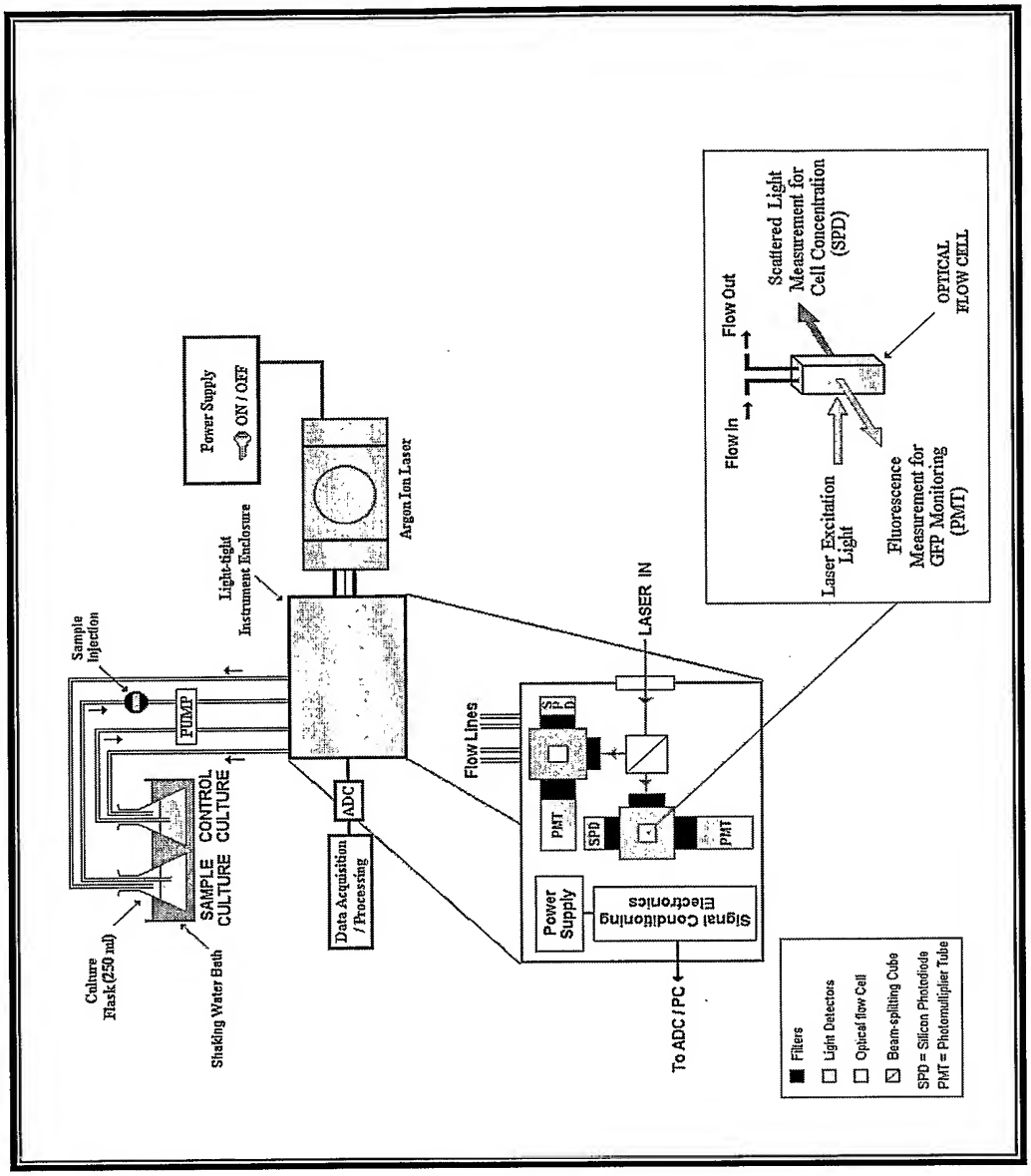

Power Supply
ON / OFF
Argon Ion Laser
Light-tight Instrument Enclosure
Sample Injection
PUMP
CONTROL CULTURE
SAMPLE CULTURE
ADC
Data Acquisition / Processing
Culture Flask (250 ml)
Shaking Water Bath
Flow Lines
LASER IN
PMT
SPD
PMT
Power Supply
Signal Conditioning Electronics
To ADC / PC
Flow Out
Scattered Light Measurement for Cell Concentration (SPD)
OPTICAL FLOW CELL
Flow In
Laser Excitation Light
Fluorescence Measurement for GFP Monitoring (PMT)
Filters
Light Detectors
Optical flow Cell
Beam-splitting Cube
SPD = Silicon Photodiode
PMT = Photomultiplier Tube

FIG. 14

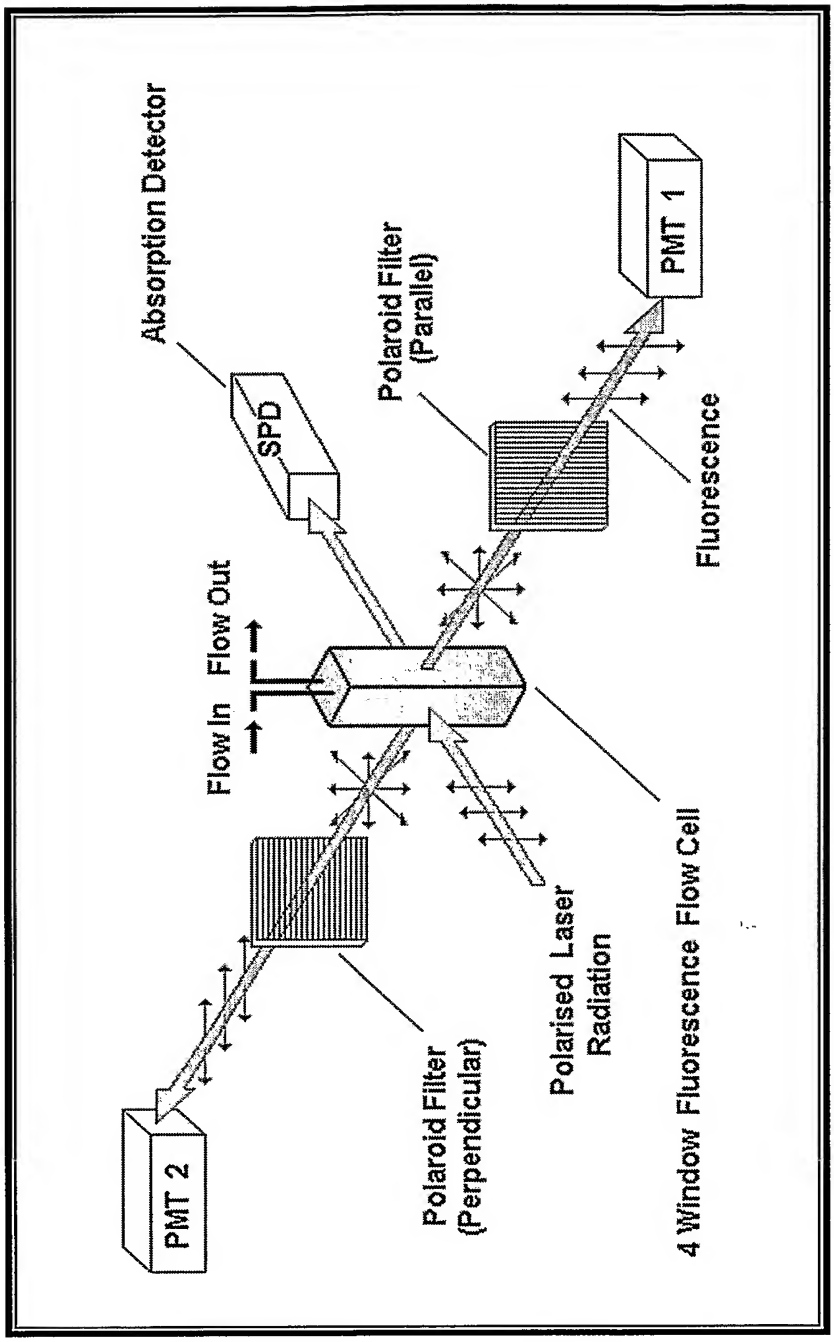

Absorption Detector
SPD
Polaroid Filter (Parallel)
PMT 1
Fluorescence
Flow In
Flow Out
4 Window Fluorescence Flow Cell
Polarised Laser Radiation
Polaroid Filter (Perpendicular)
PMT 2

FIG. 15

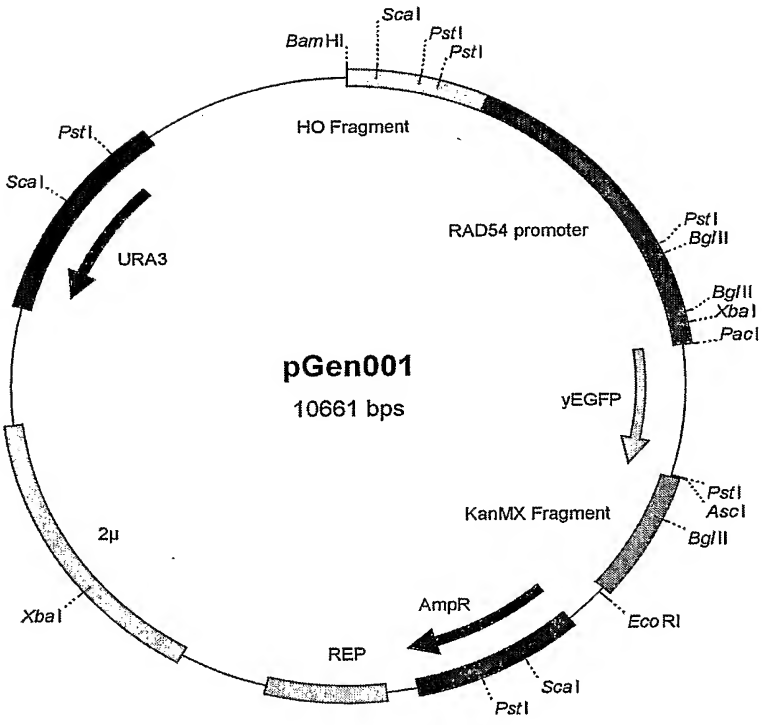

ScaI
BamHI
PstI
PstI
HO Fragment
PstI
ScaI
URA3
RAD54 promoter
PstI
BglII
BglII
XbaI
PacI
pGen001
10661 bps
yEGFP
PstI
AscI
KanMX Fragment
BglII
2µ
AmpR
EcoRI
XbaI
REP
ScaI
PstI

FIG. 16

Key:

HO sequence RAD54 Promoter *yEGFP* KanMX sequence AmpR REP 2μ sequence URA3

*GATCC*AAGCTATCTACTGAGATTTCTGGCTCTTTTGTTGTACTGTCACCTAACCACAGACCAAGCATCCAA
GCCATACTTTTTACAGCAGGAGTTACAAGGTCACTACGTCCAGTGAGAAATTTAGATAAAACACCATTTCC
TGCGAGTACTGGACCAAATCTTATGCAGCTAGAAATTCTCAATTGAGCATCAAGATAATCCAAATCTCTAA
CTTCAATGTCAAAGTTGAAATATTCTCCTTTAGAGCGCTCCATTTCTTCTATGAAGCGTTTTGCGGCAAAC
TCACCTTCAACTGTCATTGGGAATGTCTTATGATGGTTTTTTGGAATTATTATTATCCTACCATCAAGCGT
CTGACATTGCTGCAGATTTCTCCATCTCACTTTATATTTGGTGGCATTTCTACCACTTTTTTCCAACAGTG
GTTTGGTAGGGACCCTGACTGACAATTTATGACCTGCAGTACATTGTAATGCAAGACGCTGATAAACTGTT
CTACGCCTGGGATCTAACCTACCAGGTTCACCTTCAAAAGCTCTGTGTTTGGTTTTTTGCTGTATATTATA
GATTTTCTGATAGCCCTGTGTGACATTTATGACGCGGGCAGCGGAGCCATCTGCGCACATAACGTAAGAGT
TAGCCGTGACGTTTGCGATGTCTTTAATTTCACCGTTAGCCATCAGAATAGTCGTGTTTTCAGAAAGCATT
TTGATCCGACATACGATGACCTCAATGATTTAGATTATGTGTTGCACTTTTATAGACCTACCAAAAATCCA
GTGCGTACACTAATACTTTCATAAAGATACCTGAAACAATAACCAGAAAGATCGGCAAAAAAATTTTTTTT
CTTTGCCGAGATCACAAACCTACTATGACGAAAAAGCTTGAAGTTTAGATGAGTAAGGAAAATACAAGTGA
CGCTTTTATATGGTGCAAGGAACAAAAACTAAAAACAACAAGGCAAATGTGGATCTGTCATGTATGGCAAC
GACAGCAGGATGGCTCACAAAAAAAGACAAAAAAAACTAAGGCAAAAGAACAAAGCTCCTCTCCTGCTCAA
GAAACGTATTGTTGAAAAACCACCGTCGTAAGAAAGTTTTTCTGTGACCTATAATGGTTTAAAATCGGCCC
ATTTTTTTTCCCTCTTTTGTGGTCCAGTCTTTCTCATACTCGAGGGAAATTCGACACAAACAGCGGAGAAG
TGTGGCTAAACCGGCAAGTGCCTGCAAGATCCACAGAACTAACCGCACGAACTGGCGGTCAGAAAAGAGCC
TGTTCCGGAAAGAGAGAAACAGAGAAACGATCATGATGGGAAAGCGGGGATTCGGCGAAGAACGAGACTGG
AAAGGGAAAAAGAGAAATACTGGTGGAAGTATTCGGACCTTTGGCGAAGTCCGAACCCTTGAAACCCAAAG
ATGATCGATGATTCATTTTTCAATGCGCTACGGTTCCTGCCGCTCGTGGGAACCCCACGCAAAACATATTA
TTCGCTTCTCTCTGCTGACAACTCCGGTTTACGTTATACCGTATTAGGATCACTATAAGGGTTCCTTCGGG
AGGAGGGGGGAGGGGAAGAATGTACATCGTCATAAGGCCTTTATGGTGTGAAGTGGGTTTTGCGTGGAAAA
TTCGTTTTCAATGATATAGAGCCCACGCATATACGTACATACTAGTGGCCAAAAGCGTGGGGTGGGCGGAC
AAAGCTACACTGGTAAAATACAGGATTCTATGAACAATAACAACAACCAGCTCACGTTGCTGAACAGCCGA
GGTCAGCCGATGCAACCGAGGTTTCCAAAGTAGCATTTCTGTGCTAGCTATGTCTGTAGGTTTACATTTAA
TGGTGCGTGGTTCCAGCTTCATGTGCTTGCATGTGATGTCCTGCAGATGGTAAGAAGATTCTGAAAGCCGC
GCTAGGAGAAAAATATTCTGCTCGAAGATCTGTCCTCTTAAGTAGAAAGCGTGAAATTGTTGCGTTCTTGC
ATTACTACTCAACGCGTACGCAAATGCGTCTACTGCACCTGCATGATAAAGCTTATGTATCAAAAATTTAA
CATCTTGAAAATACACAAGTGGTGCAAAGATGTGTCACGTTCTGGACCTGAGTGGTGCCATGTATGCTATT
TAACATGCAAAGGGGAAGACCCTTCCGCCTTACTGCAATAATAAAAAGTATTT**TACGCGTTACCCAATATA**
GCAAAGTTTCGCGCAAAAAAAAAAATAAAAAACAATTACAAACAAAAAGAAAAAAAAGGAAATAATAGAAG
ATCTAACTGAAGCGAAGGCCAAAACTCTTCTCACTTGACGTAATAGCCGATACAAAATCTAGAGCAGCAAC
TTTTCTCTTTCTTCACTAAAGCTGCTACGAAAGTATAGAAAAATCAAACGCTCAGAACTTAGCTCTATTTC
AAGGTACCATATATATTTCCTTATAACTG*ATGTTAATTAA*CTCTAAAGGTGAAGAATTATTCACTGGTGTT
GTCCCAATTTTGGTTGAATTAGATGGTGATGTTAATGGTCACAAATTTTCTGTCTCCGGTGAAGGTGAAGG
TGATGCTACTTACGGTAAATTGACCTTAAAATTTATTTGTACTACTGGTAAATTGCCAGTTCCATGGCCAA
CCTTAGTCACTACTTTCGGTTATGGTGTTCAATGTTTTGCGAGATACCCAGATCATATGAAACAACATGAC
TTTTTCAAGTCTGCCATGCCAGAAGGTTATGTTCAAGAAAGAACTATTTTTTTCAAAGATGACGGTAACTA
CAAGACCAGAGCTGAAGTCAAGTTTGAAGGTGATACCTTAGTTAATAGAATCGAATTAAAAGGTATTGATT
TTAAAGAAGATGGTAACATTTTAGGTCACAAATTGGAATACAACTATAACTCTCACAATGTTTACATCATG
GCTGACAAACAAAAGAATGGTATCAAAGTTAACTTCAAAATTAGACACAACATTGAAGATGGTTCTGTTCA
ATTAGCTGACCATTATCAACAAAATACTCCAATTGGTGATGGTCCAGTCTTGTTACCAGACAACCATTACT
TATCCACTCAATCTGCCTTATCCAAAGATCCAAACGAAAAGAGAGACCACATGGTCTTGTTAGAATTTGTT

Figure 16 continued

*ACTGCTGCTGGTATTACCCATGGTATGGATGAATTGTACAAATAA*CTGCAGGGCGCGCCACTTCTAAATAA
GCGAATTTCTTATGATTTATGATTTTTATTATTAAATAAGTTATAAAAAAAATAAGTGTATACAAATTTTA
AAGTGACTCTTAGGTTTTAAAACGAAAATTCTTATTCTTGAGTAACTCTTTCCTGTAGGTCAGGTTGCTTT
CTCAGGTATAGTATGAGGTCGCTCTTATTGACCACACCTCTACCGGCAGATCCGCTAGGGATAACAGGGTA
ATATAGATCTGCCCGCCGGGAAGGCGAACCCGATCGGATGCATCCTCTCTGCTGCCATGATGCTGAAGTTG
TCGTTGAACATGGTTGCTGCCGGCGAGGCGGTCGAGCAGGCAGTGCAGGAGGTGTTGGACTCGGGAGTCAG
AACGGGCGACCTGCTCGGCTCGAGGTCCACTTCGGAGGTTGGCGACGCCATTGCGCTTGCAGTTAAGGAAG
CCTTGCGCAGGCAATCCGCAGCTGGTCTGAGCTAGCCTCGAGGACCCTTCTCTTTAGACTATTCTACTCTT
ATGCACGTAAAAAATTCTAGGAAATATGTATTAACTAGGAGTAAAATAACCGGCTAGTGGCATTCATATAG
CCGTCTGTTTACATCTACATCACACATTTCGAGTGTATATCTCGCAACGTTGGCGTTAAATAGGCAGTCAA
TGGCCCGACCATTCTATGGTGTTTAGGTCGATGCCATCTTTGTACGTTTAGCTTATCGATGATAAGCTGTC
AAACATGAGAATTCTTGAAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAAT
AATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCT
AAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGG
AAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTT
TGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCG
AACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACT
TTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTGTTGACGCCGGGCAAGAGCAACTCGGTCGCCGCAT
ACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAG
TAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATC
GGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGA
ACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGCAGCAATGGCAACAACGT
TGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCG
GATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGC
CGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTA
TCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTG
ATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTA
ATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGT
TCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATC
TGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCT
TTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAG
GCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCT
GCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTC
GGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTAC
AGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGG
GTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTT
TCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCA
GCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCC
CCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGA
GCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCG
GTATTTCACACCGCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATACA
CTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGA
CGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAG
GTTTTCACCGTCATCACCGAAACGCGCGAGGCAGAGCTTTGAAGAAAAATGCGCCTTATTCAATCTTTGCT
ATAAAAAATGGCCCAAAATCTCACATTGGAAGACATTTGATGACCTCATTTCTTTCAATGAAGGGCCTAAC
GGAGTTGACTAATGTTGTGGGAAATTGGAGCGATAAGCGTGCTTCTGCCGTGGCCAGGACAACGTATACTC
ATCAGATAACAGCAATACCTGATCACTACTTCGCACTAGTTTCTCGGTACTATGCATATGATCCAATATCA
AAGGAAATGATAGCATTGAAGGATGAGACTAATCCAATTGAGGAGTGGCAGCATATAGAACAGCTAAAGGG
TAGTGCTGAAGGAAGCATACGATACCCCGCATGGAATGGGATAATATCACAGGAGGTACTAGACTACCTTT

CATCCTACATAAATAGACGCATATAAGTACGCATTTAAGCATAAACACGCACTATGCCGTTCTTCTCATGT
ATATATATATACAGGCAACACGCAGATATAGGTGCGACGTGAACAGTGAGCTGTATGTGCGCAGCTCGCGT
TGCATTTTCGGAAGCGCTCGTTTTCGGAAACGCTTTGAAGTTCCTATTCCGAAGTTCCTATTCTCTAGAAA
GTATAGGAACTTCAGAGCGCTTTTGAAAACCAAAAGCGCTCTGAAGACGCACTTTCAAAAAACCAAAAACG
CACCGGACTGTAACGAGCTACTAAAATATTGCGAATACCGCTTCCACAAACATTGCTCAAAAGTATCTCTT
TGCTATATATCTCTGTGCTATATCCCTATATAACCTACCCATCCACCTTTCGCTCCTTGAACTTGCATCTA
AACTCGACCTCTACATTTTTTATGTTTATCTCTAGTATTACTCTTTAGACAAAAAAATTGTAGTAAGAACT
ATTCATAGAGTGAATCGAAAACAATACGAAAATGTAAACATTTCCTATACGTAGTATATAGAGACAAAATA
GAAGAAACCGTTCATAATTTTCTGACCAATGAAGAATCATCAACGCTATCACTTTCTGTTCACAAAGTATG
CGCAATCCACATCGGTATAGAATATAATCGGGGATGCCTTTATCTTGAAAAAATGCACCCGCAGCTTCGCT
AGTAATCAGTAAACGCGGGAAGTGGAGTCAGGCTTTTTTTATGGAAGAGAAAATAGACACCAAAGTAGCCT
TCTTCTAACCTTAACGGACCTACAGTGCAAAAAGTTATCAAGAGACTGCATTATAGAGCGCACAAAGGAGA
AAAAAAGTAATCTAAGATGCTTTGTTAGAAAAATAGCGCTCTCGGGATGCATTTTTGTAGAACAAAAAAGA
AGTATAGATTCTTTGTTGGTAAAATAGCGCTCTCGCGTTGCATTTCTGTTCTGTAAAAATGCAGCTCAGAT
TCTTTGTTTGAAAAATTAGCGCTCTCGCGTTGCATTTTTGTTTTACAAAAATGAAGCACAGATTCTTCGTT
GGTAAAATAGCGCTTTCGCGTTGCATTTCTGTTCTGTAAAAATGCAGCTCAGATTCTTTGTTTGAAAAATT
AGCGCTCTCGCGTTGCATTTTTGTTCTACAAAATGAAGCACAGATGCTTCGTTCTGCGGTAAAGCTCATCA
GCGTGGTCGTGAAGCGATTCACAGATGTCTGCCTGTTCATCCGCGTCCAGCTCGTTGAGTTTCTCCAGAAG
CGTTAATGTCTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTTTCCTGTTTGGTCACTGATGCCT
CCGTGTAAGGGGGATTTCTGTTCATGGGGGTAATGATACCGATGAAACGAGAGAGGATGCTCACGATACGG
GTTACTGATGATGAACATGCCCGGTTACTGGAACGTTGTGAGGGTAAACAACTGGCGGTATGGATGCGGCG
GGACCAGAGAAAAATCACTCAGGGTCAATGCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTA
GCCAGCAGCATCCTGCGATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTTCCAGACTT
TACGAAACACGGAAACCGAAGACCATTCATGTTGTTGCTCAGGTCGCAGACGTTTTGCAGCAGCAGTCGCT
TCACGTTCGCTCGCGTATCGGTGATTCATTCTGCTAACCAGTAAGGCAACCCCGCCAGCCTAGCCGGGTCC
TCAACGACAGGAGCACGATCATGCGCACCCGTGGCCAGGACCCAACGCTGCGGGGGGGGGGGGGGTTTTCT
TTCCAATTTTTTTTTTTTCGTCATTATAGAAATCATTACGACCGAGATTCCCGGGTAATAACTGATATAAT
TAAATTGAAGCTCTAATTTGTGAGTTTAGTATACATGCATTTACTTATAATACAGTTTTTTAGTTTTGCTG
GCCGCATCTTCTCAAATATGCTTCCCAGCCTGCTTTTCTGTAACGTTCACCCTCTACCTTAGCATCCCTTC
CCTTTGCAAATAGTCCTCTTCCAACAATAATAATGTCAGATCCTGTAGAGACCACATCATCCACGGTTCTA
TACTGTTGACCCAATGCGTCTCCCTTGTCATCTAAACCCACACCGGGGTGTCATAATCAACCAATCGTAACC
TTCATCTCTTCCACCCATGTCTCTTTGAGCAATAAAGCCGATAACAAAATCTTTGTCGCTCTTCGCAATGT
CAACAGTACCCTTAGTATATTCTCCAGTAGCTAGGGAGCCCTTGCATGACAATTCTGCTAACATCAAAAGG
CCTCTAGGTTCCTTTGTTACTTCTTCCGCCGCCTGCTTCAAACCGCTAACAATACCTGGGCCCACCACACC
GTGTGCATTCGTAATGTCTGCCCATTCTGCTATTCTGTATACACCCGCAGAGTACTGCAATTTGACTGTAT
TACCAATGTCAGCAAATTTTCTGTCTTCGAAGAGTAAAAAATTGTACTTGGCGGATAATGCCTTTAGCGGC
TTAACTGTGCCCTCCATGGAAAAATCAGTCAAGATATCCACATGTGTTTTTAGTAAACAAATTTTGGGACC
TAATGCTTCAACTAACTCCAGTAATTCCTTGGTGGTACGAACATCCAATGAAGCACACAAGTTTGTTTGCT
TTTCGTGCATGATATTAAATAGCTTGGCAGCAACAGGACTAGGATGAGTAGCAGCACGTTCCTTATATGTA
GCTTTCGACATGATTTATCTTCGTTTCCTGCAGGTTTTTGTTCTGTGCAGTTGGGTTAAGAATACTGGGCA
ATTTCATGTTTCTTCAACACCACATATGCGTATATATACCAATCTAAGTCTGTGCTCCTTCCTTCGTTCTT
CCTTCTGCTCGGAGATTACCGAATCAAAAAAATTTCAAAGAAACCGGAATCAAAAAAAGAACAAAAAAAA
AAAAGATGAATTGAAACCCCCCCCCCCCCCGATGCGCCGCGTGCGGCTGCTGGAGATGGCGGACGCGATGG
ATATGTTCTGCCAAGGGTTGGTTTGCGCATTCACAGTTCTCCGCAAGAATTGATTGGCTCCAATTCTTGGA
GTGGTGAATCCGTTAGCGAGGTGCCGCCGGCTTCCATTCAGGTCGAGGTGGCCCGGCTCCATGCACCGCGA
CGCAACGCGGGGAGGCAGACAAGGTATAGGGCGGCGCCTACAATCCATGCCAACCCGTTCCATGTGCTCGC
CGAGGCGGCATAAATCGCCGTGACGATCAGCGGTCCAGTGATCGAAGTTAGGCTGGTAAGAGCCGCGAGCG
ATCCTTGAAGCTGTCCCTGATGGTCGTCATCTACCTGCCTGGACAGCATGGCCTGCAACGCGGGCATCCCG
ATGCCGCCGGAAGCGAGAAGAATCATAATGGGGAAGGCCATCCAGCCTCGCGTCGCGAACGCCAGCAAGAC
GTAGCCCAGCGCGTCGGCCGCCATGCCGGCGATAATGGCCTGCTTCTCGCCGAAACGTTTGGTGGCGGGAC
CAGTGACGAAGGCTTGAGCGAGGGCGTGCAAGATTCCGAATACCGCAAGCGACAGGCCGATCATCGTCGCG
CTCCAGCGAAAGCGGTCCTCGCCGAAAATGACCCAGAGCGCTGCCGGCACCTGTCCTACGAGTTGCATGAT

AAAGAAGACAGTCATAAGTGCGGCGACGATAGTCATGCCCCGCGCCCACCGGAAGGAGCTGACTGGGTTGA
AGGCTCTCAAGGGCATCGGTCGACGCTCTCCCTTATGCGACTCCTGCATTAGGAAGCAGCCCAGTAGTAGG
TTGAGGCCGTTGAGCACCGCCGCCGCAAGGAATGGTGCATGCAAGGAGATGGCGCCCAACAGTCCCCCGGC
CACGGGGCCTGCCACCATACCCACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCCC
CATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATGCGT
CCGGCGTAGAG

FIG. 17

```
   1  GATCCAAGCT ATCTACTGAG ATTTCTGGCT CTTTTGTTGT ACTGTCACCT
  51  AACCACAGAC CAAGCATCCA AGCCATACTT TTTACAGCAG GAGTTACAAG
 101  GTCACTACGT CCAGTGAGAA ATTTAGATAA AACACCATTT CCTGCGAGTA
 151  CTGGACCAAA TCTTATGCAG CTAGAAATTC TCAATTGAGC ATCAAGATAA
 201  TCCAAATCTC TAACTTCAAT GTCAAAGTTG AAATATTCTC CTTTAGAGCG
 251  CTCCATTTCT TCTATGAAGC GTTTTGCGGC AAACTCACCT TCAACTGTCA
 301  TTGGGAATGT CTTATGATGG TTTTTTGGAA TTATTATTAT CCTACCATCA
 351  AGCGTCTGAC ATTGCTGCAG ATTTCTCCAT CTCACTTTAT ATTTGGTGGC
 401  ATTTCTACCA CTTTTTTCCA ACAGTGGTTT GGTAGGGACC CTGACTGACA
 451  ATTTATGACC TGCAGTACAT TGTAATGCAA GACGCTGATA AACTGTTCTA
 501  CGCCTGGGAT CTAACCTACC AGGTTCACCT TCAAAAGCTC TGTGTTTGGT
 551  TTTTTGCTGT ATATTATAGA TTTTCTGATA GCCCTGTGTG ACATTTATGA
 601  CGCGGGCAGC GGAGCCATCT GCGCACATAA CGTAAGAGTT AGCCGTGACG
 651  TTTGCGATGT CTTTAATTTC ACCGTTAGCC ATCAGAATAG TCGTGTTTTC
 701  AGAAAGCATT TTGATCCGAC ATACGATGAC CTCAATGATT TAGATTATGT
 751  GTTGCACTTT TATAGACCTA CCAAAAATCC AGTGCGTACA CTAATACTTT
 801  CATAAAGATA CCTGAAACAA TAACCAGAAA GATCGGCAAA AAAATTTTTT
 851  TTCTTTGCCG AGATCACAAA CCTACTATGA CGAAAAAGCT TGAAGTTTAG
 901  ATGAGTAAGG AAAATACAAG TGACGCTTTT ATATGGTGCA AGGAACAAAA
 951  ACTAAAAACA ACAAGGCAAA TGTGGATCTG TCATGTATGG CAACGACAGC
1001  AGGATGGCTC ACAAAAAAAG ACAAAAAAAA CTAAGGCAAA AGAACAAAGC
1051  TCCTCTCCTG CTCAAGAAAC GTATTGTTGA AAAACCACCG TCGTAAGAAA
1101  GTTTTTCTGT GACCTATAAT GGTTTAAAAT CGGCCCATTT TTTTTCCCTC
1151  TTTTGTGGTC CAGTCTTTCT CATACTCGAG GGAAATTCGA CACAAACAGC
1201  GGAGAAGTGT GGCTAAACCG GCAAGTGCCT GCAAGATCCA CAGAACTAAC
1251  CGCACGAACT GGCGGTCAGA AAAGAGCCTG TTCCGGAAAG AGAGAAACAG
1301  AGAAACGATC ATGATGGGAA AGCGGGGATT CGGCGAAGAA CGAGACTGGA
1351  AAGGGAAAAA GAGAAATACT GGTGGAAGTA TTCGGACCTT TGGCGAAGTC
1401  CGAACCCTTG AAACCCAAAG ATGATCGATG ATTCATTTTT CAATGCGCTA
1451  CGGTTCCTGC CGCTCGTGGG AACCCCACGC AAAACATATT ATTCGCTTCT
1501  CTCTGCTGAC AACTCCGGTT TACGTTATAC CGTATTAGGA TCACTATAAG
1551  GGTTCCTTCG GGAGGAGGGG GGAGGGGAAG AATGTACATC GTCATAAGGC
1601  CTTTATGGTG TGAAGTGGGT TTTGCGTGGA AAATTCGTTT TCAATGATAT
1651  AGAGCCCACG CATATACGTA CATACTAGTG GCCAAAAGCG TGGGGTGGGC
1701  GGACAAAGCT ACACTGGTAA AATACAGGAT TCTATGAACA ATAACAACAA
1751  CCAGCTCACG TTGCTGAACA GCCGAGGTCA GCCGATGCAA CCGAGGTTTC
1801  CAAAGTAGCA TTTCTGTGCT AGCTATGTCT GTAGGTTTAC ATTTAATGGT
1851  GCGTGGTTCC AGCTTCATGT GCTTGCATGT GATGTCCTGC AGATGGTAAG
1901  AAGATTCTGA AAGCCGCGCT AGGAGAAAAA TATTCTGCTC GAAGATCTGT
1951  CCTCTTAAGT AGAAAGCGTG AAATTGTTGC GTTCTTGCAT TACTACTCAA
2001  CGCGTACGCA AATGCGTCTA CTGCACCTGC ATGATAAAGC TTATGTATCA
2051  AAAATTTAAC ATCTTGAAAA TACACAAGTG GTGCAAAGAT GTGTCACGTT
2101  CTGGACCTGA GTGGTGCCAT GTATGCTATT TAACATGCAA AGGGGAAGAC
2151  CCTTCCGCCT TACTGCAATA ATAAAAAGTA TTTTACGCGT TACCCAATAT
2201  AGCAAAGTTT CGCGCAAAAA AAAAAATAAA AAACAATTAC AAACAAAAAG
2251  AAAAAAAAGG AAATAATAGA AGATCTAACT GAAGCGAAGG CCAAAACTCT
2301  TCTCACTTGA CGTAATAGCC GATACAAAAT CTAGAGCAGC AACTTTTCTC
2351  TTTCTTCACT AAAGCTGCTA CGAAAGTATA GAAAAATCAA ACGCTCAGAA
2401  CTTAGCTCTA TTTCAAGGTA CCATATATAT TTCCTTATAA CTGATGTTAA
2451  TTAACTCTAA AGGTGAAGAA TTATTCACTG GTGTTGTCCC AATTTTGGTT
2501  GAATTAGATG GTGATGTTAA TGGTCACAAA TTTTCTGTCT CCGGTGAAGG
```

Fig 17 continued

```
2551  TGAAGGTGAT GCTACTTACG GTAAATTGAC CTTAAAATTT ATTTGTACTA
2601  CTGGTAAATT GCCAGTTCCA TGGCCAACCT TAGTCACTAC TTTCGGTTAT
2651  GGTGTTCAAT GTTTTGCGAG ATACCCAGAT CATATGAAAC AACATGACTT
2701  TTTCAAGTCT GCCATGCCAG AAGGTTATGT TCAAGAAAGA ACTATTTTTT
2751  TCAAAGATGA CGGTAACTAC AAGACCAGAG CTGAAGTCAA GTTTGAAGGT
2801  GATACCTTAG TTAATAGAAT CGAATTAAAA GGTATTGATT TTAAAGAAGA
2851  TGGTAACATT TTAGGTCACA AATTGGAATA CAACTATAAC TCTCACAATG
2901  TTTACATCAT GGCTGACAAA CAAAAGAATG GTATCAAAGT TAACTTCAAA
2951  ATTAGACACA ACATTGAAGA TGGTTCTGTT CAATTAGCTG ACCATTATCA
3001  ACAAAATACT CCAATTGGTG ATGGTCCAGT CTTGTTACCA GACAACCATT
3051  ACTTATCCAC TCAATCTGCC TTATCCAAAG ATCCAAACGA AAAGAGAGAC
3101  CACATGGTCT TGTTAGAATT TGTTACTGCT GCTGGTATTA CCCATGGTAT
3151  GGATGAATTG TACAAATAAC TGCAGGGCGC GCCACTTCTA AATAAGCGAA
3201  TTTCTTATGA TTTATGATTT TTATTATTAA ATAAGTTATA AAAAAAATAA
3251  GTGTATACAA ATTTTAAAGT GACTCTTAGG TTTTAAAACG AAAATTCTTA
3301  TTCTTGAGTA ACTCTTTCCT GTAGGTCAGG TTGCTTTCTC AGGTATAGTA
3351  TGAGGTCGCT CTTATTGACC ACACCTCTAC CGGCAGATCC GCTAGGGATA
3401  ACAGGGTAAT ATAGATCTGC CCGCCGGGAA GGCGAACCCG ATCGGATGCA
3451  TCCTCTCTGC TGCCATGATG CTGAAGTTGT CGTTGAACAT GGTTGCTGCC
3501  GGCGAGGCGG TCGAGCAGGC AGTGCAGGAG GTGTTGGACT CGGGAGTCAG
3551  AACGGGCGAC CTGCTCGGCT CGAGCTCCAC TTCGGAGGTT GGCGACGCCA
3601  TTGCGCTTGC AGTTAAGGAA GCCTTGCGCA GGCAATCCGC AGCTGGTCTG
3651  AGCTAGCCTC GAGGACCCTT CTCTTTAGAC TATTCTACTC TTATGCACGT
3701  AAAAAATTCT AGGAAATATG TATTAACTAG GAGTAAAATA ACCGGCTAGT
3751  GGCATTCATA TAGCCGTCTG TTTACATCTA CATCACACAT TTCGAGTGTA
3801  TATCTCGCAA CGTTGGCGTT AAATAGGCAG TCAATGGCCC GACCATTCTA
3851  TGGTGTTTAG GTCGATGCCA TCTTTGTACG TTTAGCTTAT CGATGATAAG
3901  CTGTCAAACA TGAGAATTCT TGAAGACGAA AGGGCCTCGT GATACGCCTA
3951  TTTTTATAGG TTAATGTCAT GATAATAATG GTTTCTTAGA CGTCAGGTGG
4001  CACTTTTCGG GGAAATGTGC GCGGAACCCC TATTTGTTTA TTTTTCTAAA
4051  TACATTCAAA TATGTATCCG CTCATGAGAC AATAACCCTG ATAAATGCTT
4101  CAATAATATT GAAAAAGGAA GAGTATGAGT ATTCAACATT TCCGTGTCGC
4151  CCTTATTCCC TTTTTTGCGG CATTTTGCCT TCCTGTTTTT GCTCACCCAG
4201  AAACGCTGGT GAAAGTAAAA GATGCTGAAG ATCAGTTGGG TGCACGAGTG
4251  GGTTACATCG AACTGGATCT CAACAGCGGT AAGATCCTTG AGAGTTTTCG
4301  CCCCGAAGAA CGTTTTCCAA TGATGAGCAC TTTTAAAGTT CTGCTATGTG
4351  GCGCGGTATT ATCCCGTGTT GACGCCGGGC AAGAGCAACT CGGTCGCCGC
4401  ATACACTATT CTCAGAATGA CTTGGTTGAG TACTCACCAG TCACAGAAAA
4451  GCATCTTACG GATGGCATGA CAGTAAGAGA ATTATGCAGT GCTGCCATAA
4501  CCATGAGTGA TAACACTGCG GCCAACTTAC TTCTGACAAC GATCGGAGGA
4551  CCGAAGGAGC TAACCGCTTT TTTGCACAAC ATGGGGGATC ATGTAACTCG
4601  CCTTGATCGT TGGGAACCGG AGCTGAATGA AGCCATACCA AACGACGAGC
4651  GTGACACCAC GATGCCTGCA GCAATGGCAA CAACGTTGCG CAAACTATTA
4701  ACTGGCGAAC TACTTACTCT AGCTTCCCGG CAACAATTAA TAGACTGGAT
4751  GGAGGCGGAT AAAGTTGCAG GACCACTTCT GCGCTCGGCC CTTCCGGCTG
4801  GCTGGTTTAT TGCTGATAAA TCTGGAGCCG GTGAGCGTGG GTCTCGCGGT
4851  ATCATTGCAG CACTGGGGCC AGATGGTAAG CCCTCCCGTA TCGTAGTTAT
4901  CTACACGACG GGGAGTCAGG CAACTATGGA TGAACGAAAT AGACAGATCG
4951  CTGAGATAGG TGCCTCACTG ATTAAGCATT GGTAACTGTC AGACCAAGTT
5001  TACTCATATA TACTTTAGAT TGATTTAAAA CTTCATTTTT AATTTAAAAG
5051  GATCTAGGTG AAGATCCTTT TTGATAATCT CATGACCAAA ATCCCTTAAC
5101  GTGAGTTTTC GTTCCACTGA GCGTCAGACC CCGTAGAAAA GATCAAAGGA
5151  TCTTCTTGAG ATCCTTTTTT TCTGCGCGTA ATCTGCTGCT TGCAAACAAA
5201  AAAACCACCG CTACCAGCGG TGGTTTGTTT GCCGGATCAA GAGCTACCAA
```

Fig 17 continued

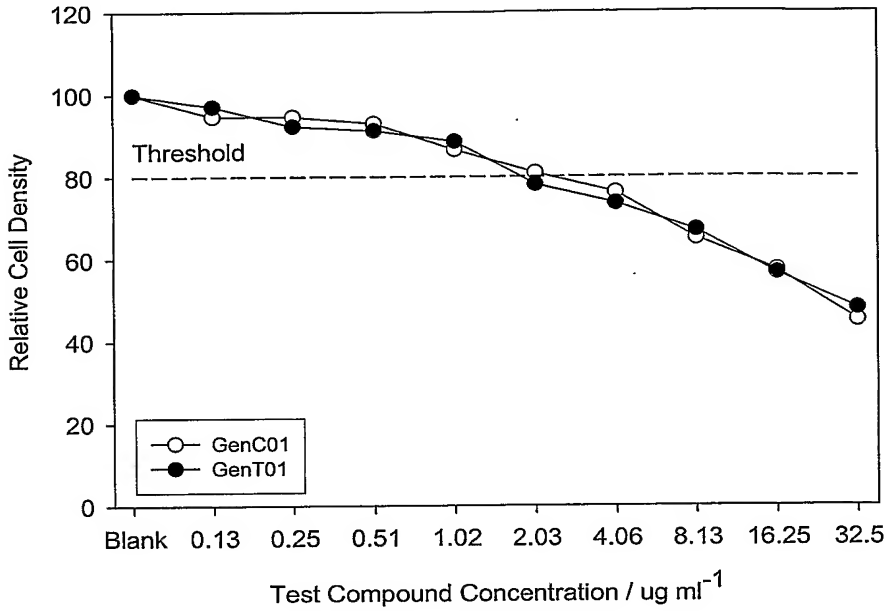
```
5251  CTCTTTTTCC GAAGGTAACT GGCTTCAGCA GAGCGCAGAT ACCAAATACT
5301  GTCCTTCTAG TGTAGCCGTA GTTAGGCCAC CACTTCAAGA ACTCTGTAGC
5351  ACCGCCTACA TACCTCGCTC TGCTAATCCT GTTACCAGTG GCTGCTGCCA
5401  GTGGCGATAA GTCGTGTCTT ACCGGGTTGG ACTCAAGACG ATAGTTACCG
5451  GATAAGGCGC AGCGGTCGGG CTGAACGGGG GGTTCGTGCA CACAGCCCAG
5501  CTTGGAGCGA ACGACCTACA CCGAACTGAG ATACCTACAG CGTGAGCTAT
5551  GAGAAAGCGC CACGCTTCCC GAAGGGAGAA AGGCGGACAG GTATCCGGTA
5601  AGCGGCAGGG TCGGAACAGG AGAGCGCACG AGGGAGCTTC CAGGGGGAAA
5651  CGCCTGGTAT CTTTATAGTC CTGTCGGGTT TCGCCACCTC TGACTTGAGC
5701  GTCGATTTTT GTGATGCTCG TCAGGGGGGC GGAGCCTATG GAAAAACGCC
5751  AGCAACGCGG CCTTTTTACG GTTCCTGGCC TTTTGCTGGC CTTTTGCTCA
5801  CATGTTCTTT CCTGCGTTAT CCCCTGATTC TGTGGATAAC CGTATTACCG
5851  CCTTTGAGTG AGCTGATACC GCTCGCCGCA GCCGAACGAC CGAGCGCAGC
5901  GAGTCAGTGA GCGAGGAAGC GGAAGAGCGC CTGATGCGGT ATTTTCTCCT
5951  TACGCATCTG TGCGGTATTT CACACCGCAT ATGGTGCACT CTCAGTACAA
6001  TCTGCTCTGA TGCCGCATAG TTAAGCCAGT ATACACTCCG CTATCGCTAC
6051  GTGACTGGGT CATGGCTGCG CCCCGACACC CGCCAACACC CGCTGACGCG
6101  CCCTGACGGG CTTGTCTGCT CCCGGCATCC GCTTACAGAC AAGCTGTGAC
6151  CGTCTCCGGG AGCTGCATGT GTCAGAGGTT TTCACCGTCA TCACCGAAAC
6201  GCGCGAGGCA GAGCTTTGAA GAAAAATGCG CCTTATTCAA TCTTTGCTAT
6251  AAAAAATGGC CCAAAATCTC ACATTGGAAG ACATTTGATG ACCTCATTTC
6301  TTTCAATGAA GGGCCTAACG GAGTTGACTA ATGTTGTGGG AAATTGGAGC
6351  GATAAGCGTG CTTCTGCCGT GGCCAGGACA ACGTATACTC ATCAGATAAC
6401  AGCAATACCT GATCACTACT TCGCACTAGT TTCTCGGTAC TATGCATATG
6451  ATCCAATATC AAAGGAAATG ATAGCATTGA AGGATGAGAC TAATCCAATT
6501  GAGGAGTGGC AGCATATAGA ACAGCTAAAG GGTAGTGCTG AAGGAAGCAT
6551  ACGATACCCC GCATGGAATG GGATAATATC ACAGGAGGTA CTAGACTACC
6601  TTTCATCCTA CATAAATAGA CGCATATAAG TACGCATTTA AGCATAAACA
6651  CGCACTATGC CGTTCTTCTC ATGTATATAT ATATACAGGC AACACGCAGA
6701  TATAGGTGCG ACGTGAACAG TGAGCTGTAT GTGCGCAGCT CGCGTTGCAT
6751  TTTCGGAAGC GCTCGTTTTC GGAAACGCTT TGAAGTTCCT ATTCCGAAGT
6801  TCCTATTCTC TAGAAAGTAT AGGAACTTCA GAGCGCTTTT GAAAACCAAA
6851  AGCGCTCTGA AGACGCACTT TCAAAAAACC AAAAACGCAC CGGACTGTAA
6901  CGAGCTACTA AAATATTGCG AATACCGCTT CCACAAACAT TGCTCAAAAG
6951  TATCTCTTTG CTATATATCT CTGTGCTATA TCCCTATATA ACCTACCCAT
7001  CCACCTTTCG CTCCTTGAAC TTGCATCTAA ACTCGACCTC TACATTTTTT
7051  ATGTTTATCT CTAGTATTAC TCTTTAGACA AAAAAATTGT AGTAAGAACT
7101  ATTCATAGAG TGAATCGAAA ACAATACGAA AATGTAAACA TTTCCTATAC
7151  GTAGTATATA GAGACAAAAT AGAAGAAACC GTTCATAATT TTCTGACCAA
7201  TGAAGAATCA TCAACGCTAT CACTTTCTGT TCACAAAGTA TGCGCAATCC
7251  ACATCGGTAT AGAATATAAT CGGGGATGCC TTTATCTTGA AAAAATGCAC
7301  CCGCAGCTTC GCTAGTAATC AGTAAACGCG GGAAGTGGAG TCAGGCTTTT
7351  TTTATGGAAG AGAAAATAGA CACCAAAGTA GCCTTCTTCT AACCTTAACG
7401  GACCTACAGT GCAAAAAGTT ATCAAGAGAC TGCATTATAG AGCGCACAAA
7451  GGAGAAAAAA AGTAATCTAA GATGCTTTGT TAGAAAAATA GCGCTCTCGG
7501  GATGCATTTT TGTAGAACAA AAAAGAAGTA TAGATTCTTT GTTGGTAAAA
7551  TAGCGCTCTC GCGTTGCATT TCTGTTCTGT AAAAATGCAG CTCAGATTCT
7601  TTGTTTGAAA AATTAGCGCT CTCGCGTTGC ATTTTTGTTT TACAAAAATG
7651  AAGCACAGAT TCTTCGTTGG TAAAATAGCG CTTTCGCGTT GCATTTCTGT
7701  TCTGTAAAAA TGCAGCTCAG ATTCTTTGTT TGAAAAATTA GCGCTCTCGC
7751  GTTGCATTTT TGTTCTACAA AATGAAGCAC AGATGCTTCG TTCTGCGGTA
7801  AAGCTCATCA GCGTGGTCGT GAAGCGATTC ACAGATGTCT GCCTGTTCAT
7851  CCGCGTCCAG CTCGTTGAGT TTCTCCAGAA GCGTTAATGT CTGGCTTCTG
7901  ATAAAGCGGG CCATGTTAAG GGCGGTTTTT TCCTGTTTGG TCACTGATGC
```

Figure 17 continued


```
 7951  CTCCGTGTAA GGGGGATTTC TGTTCATGGG GGTAATGATA CCGATGAAAC
 8001  GAGAGAGGAT GCTCACGATA CGGGTTACTG ATGATGAACA TGCCCGGTTA
 8051  CTGGAACGTT GTGAGGGTAA ACAACTGGCG GTATGGATGC GGCGGGACCA
 8101  GAGAAAAATC ACTCAGGGTC AATGCCAGCG CTTCGTTAAT ACAGATGTAG
 8151  GTGTTCCACA GGGTAGCCAG CAGCATCCTG CGATGCAGAT CCGGAACATA
 8201  ATGGTGCAGG GCGCTGACTT CCGCGTTTCC AGACTTTACG AAACACGGAA
 8251  ACCGAAGACC ATTCATGTTG TTGCTCAGGT CGCAGACGTT TTGCAGCAGC
 8301  AGTCGCTTCA CGTTCGCTCG CGTATCGGTG ATTCATTCTG CTAACCAGTA
 8351  AGGCAACCCC GCCAGCCTAG CCGGGTCCTC AACGACAGGA GCACGATCAT
 8401  GCGCACCCGT GGCCAGGACC CAACGCTGCG GGGGGGGGGG GGGTTTTCTT
 8451  TCCAATTTTT TTTTTTTCGT CATTATAGAA ATCATTACGA CCGAGATTCC
 8501  CGGGTAATAA CTGATATAAT TAAATTGAAG CTCTAATTTG TGAGTTTAGT
 8551  ATACATGCAT TTACTTATAA TACAGTTTTT TAGTTTTGCT GGCCGCATCT
 8601  TCTCAAATAT GCTTCCCAGC CTGCTTTTCT GTAACGTTCA CCCTCTACCT
 8651  TAGCATCCCT TCCCTTTGCA AATAGTCCTC TTCCAACAAT AATAATGTCA
 8701  GATCCTGTAG AGACCACATC ATCCACGGTT CTATACTGTT GACCCAATGC
 8751  GTCTCCCTTG TCATCTAAAC CCACACCGGG TGTCATAATC AACCAATCGT
 8801  AACCTTCATC TCTTCCACCC ATGTCTCTTT GAGCAATAAA GCCGATAACA
 8851  AAATCTTTGT CGCTCTTCGC AATGTCAACA GTACCCTTAG TATATTCTCC
 8901  AGTAGCTAGG GAGCCCTTGC ATGACAATTC TGCTAACATC AAAAGGCCTC
 8951  TAGGTTCCTT TGTTACTTCT TCCGCCGCCT GCTTCAAACC GCTAACAATA
 9001  CCTGGGCCCA CCACACCGTG TGCATTCGTA ATGTCTGCCC ATTCTGCTAT
 9051  TCTGTATACA CCCGCAGAGT ACTGCAATTT GACTGTATTA CCAATGTCAG
 9101  CAAATTTTCT GTCTTCGAAG AGTAAAAAAT TGTACTTGGC GGATAATGCC
 9151  TTTAGCGGCT TAACTGTGCC CTCCATGGAA AAATCAGTCA AGATATCCAC
 9201  ATGTGTTTTT AGTAAACAAA TTTTGGGACC TAATGCTTCA ACTAACTCCA
 9251  GTAATTCCTT GGTGGTACGA ACATCCAATG AAGCACACAA GTTTGTTTGC
 9301  TTTTCGTGCA TGATATTAAA TAGCTTGGCA GCAACAGGAC TAGGATGAGT
 9351  AGCAGCACGT TCCTTATATG TAGCTTTCGA CATGATTTAT CTTCGTTTCC
 9401  TGCAGGTTTT TGTTCTGTGC AGTTGGGTTA AGAATACTGG GCAATTTCAT
 9451  GTTTCTTCAA CACCACATAT GCGTATATAT ACCAATCTAA GTCTGTGCTC
 9501  CTTCCTTCGT TCTTCCTTCT GCTCGGAGAT TACCGAATCA AAAAAATTTC
 9551  AAAGAAACCG GAATCAAAAA AAAGAACAAA AAAAAAAAAG ATGAATTGAA
 9601  ACCCCCCCCC CCCCCGATGC GCCGCGTGCG GCTGCTGGAG ATGGCGGACG
 9651  CGATGGATAT GTTCTGCCAA GGGTTGGTTT GCGCATTCAC AGTTCTCCGC
 9701  AAGAATTGAT TGGCTCCAAT TCTTGGAGTG GTGAATCCGT TAGCGAGGTG
 9751  CCGCCGGCTT CCATTCAGGT CGAGGTGGCC CGGCTCCATG CACCGCGACG
 9801  CAACGCGGGG AGGCAGACAA GGTATAGGGC GGCGCCTACA ATCCATGCCA
 9851  ACCCGTTCCA TGTGCTCGCC GAGGCGGCAT AAATCGCCGT GACGATCAGC
 9901  GGTCCAGTGA TCGAAGTTAG GCTGGTAAGA GCCGCGAGCG ATCCTTGAAG
 9951  CTGTCCCTGA TGGTCGTCAT CTACCTGCCT GGACAGCATG GCCTGCAACG
10001  CGGGCATCCC GATGCCGCCG GAAGCGAGAA GAATCATAAT GGGGAAGGCC
10051  ATCCAGCCTC GCGTCGCGAA CGCCAGCAAG ACGTAGCCCA GCGCGTCGGC
10101  CGCCATGCCG GCGATAATGG CCTGCTTCTC GCCGAAACGT TTGGTGGCGG
10151  GACCAGTGAC GAAGGCTTGA GCGAGGGCGT GCAAGATTCC GAATACCGCA
10201  AGCGACAGGC CGATCATCGT CGCGCTCCAG CGAAAGCGGT CCTCGCCGAA
10251  AATGACCCAG AGCGCTGCCG GCACCTGTCC TACGAGTTGC ATGATAAAGA
10301  AGACAGTCAT AAGTGCGGCG ACGATAGTCA TGCCCCGCGC CCACCGGAAG
10351  GAGCTGACTG GGTTGAAGGC TCTCAAGGGC ATCGGTCGAC GCTCTCCCTT
10401  ATGCGACTCC TGCATTAGGA AGCAGCCCAG TAGTAGGTTG AGGCCGTTGA
10451  GCACCGCCGC CGCAAGGAAT GGTGCATGCA AGGAGATGGC GCCCAACAGT
10501  CCCCCGGCCA CGGGGCCTGC CACCATACCC ACGCCGAAAC AAGCGCTCAT
10551  GAGCCCGAAG TGGCGAGCCC GATCTTCCCC ATCGGTGATG TCGGCGATAT
10601  AGGCGCCAGC AACCGCACCT GTGGCGCCGG TGATGCCGGC CACGATGCGT
10651  CCGGCGTAGA G
```

FIG. 18

Cytotoxicity Profile:



Genotoxicity Profile:

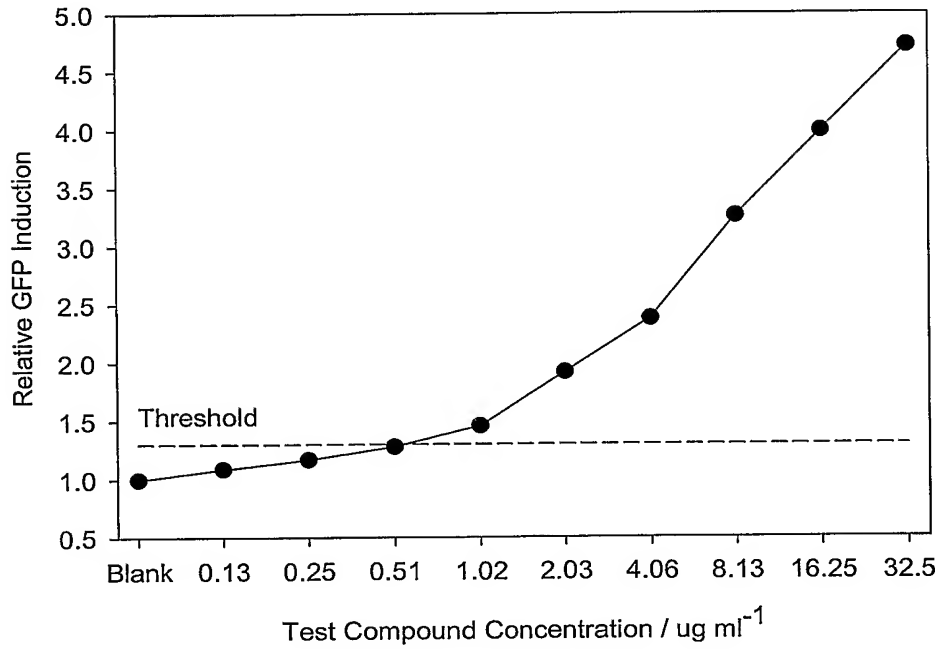
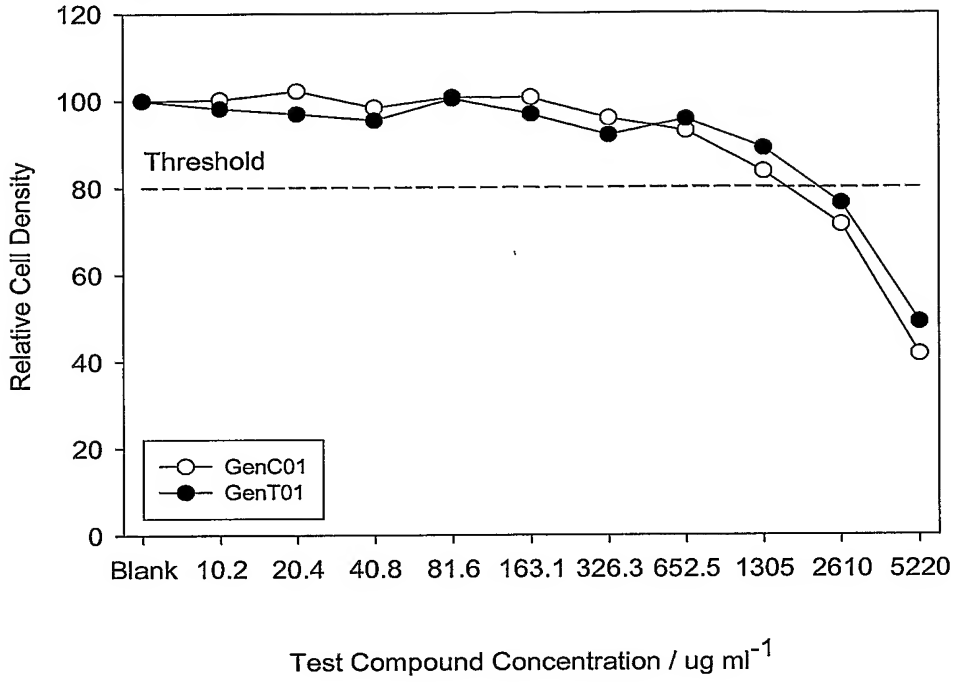


FIG. 19

Cytotoxicity Profile:



Genotoxicity Profile:

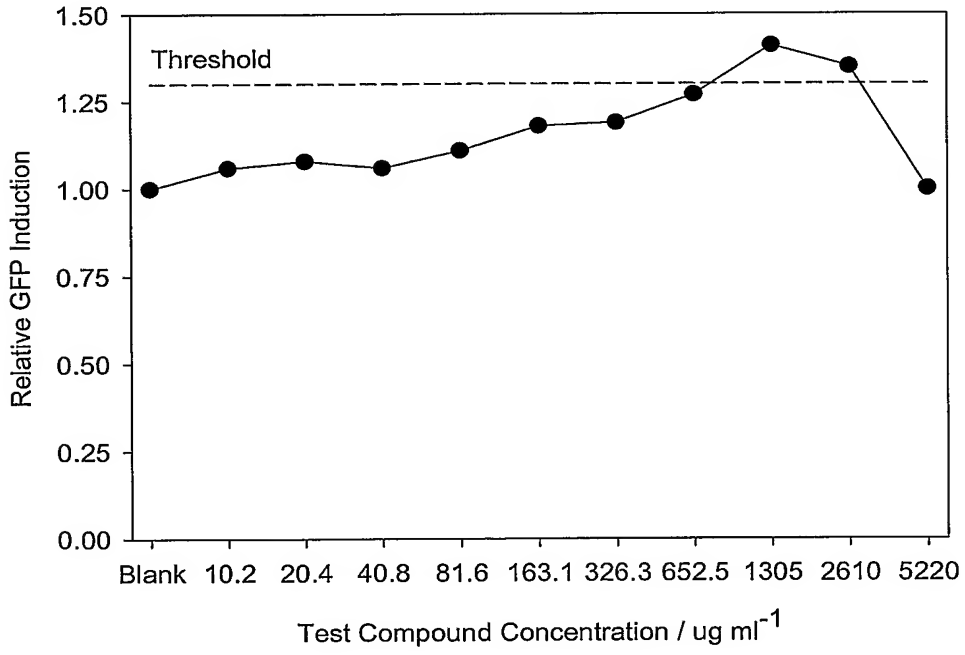


Fig.20 [1/3]

GREEN SCREEN RESULTS

ALTERNATIVE TEST DATA

TEST COMPOUND	CAS No.	GenC01 Strain Cytotoxicity	LEC / µg ml^{-1}	Test Range µg ml^{-1}	Test Range mM	GenT01 Strain Genotoxicity	LEC / µg ml^{-1}	Test Range µg ml^{-1}	Test Range mM	FP	Carcinogenicity	Ames Test	ML Test	UmuC Test	MNT in vitro	MNT in vivo	Chrom Abs
2-Acetamidofluorene	53-96-3	++	28.8	115	0.52	-		115	0.52		+	+/- (MA)	+	+ (MA)	+	+	+
Acetylsalicylic Acid	50-78-2	-		867	4.81	-		867	4.81			-			-	-	+
Actinomycin D	50-76-0	+	62.5	500	0.40	-		500	0.40		+	-	+	-		+/-	+
Acycloguanosine	59277-89-3	+	56.3	112.5	0.50	-		112.5	0.50		-						
9-Aminoacridine	90-45-9	++	8	128	0.55	-		128	0.55			+/-	+	+			+ (p)
2-Aminoanthracene	613-13-8	++	14.25	57	0.29	-		57	0.29		+	+ (MA)	+ (MA)	+ (MA)	+		
2-Amino-4-nitrophenol	99-57-0	++	19.25	77	0.50	++	9.6	77	0.50		+/-	+/- (MA)	+			-	+
4-Aminophenol	123-30-8	-		504	4.62	-		504	4.62			-	+			+	
3-Amino-1,2,4-Triazole	61-82-5	+	840	840	9.99	+	840	840	9.99		+/-	-	-				-
Ampicillin (Na salt)	69-52-3	-		8000	22.9	-		8000	22.9		-	-	-			-	
Aniline	62-53-3	++	2553	20420	219.3	-		20420	219.3		+/-	-	+		-	+	+
o-Anisidine	90-04-0	++	14.2	437	3.55	-		437	3.55		+	+/- (MA)	+	-			+
Aphidicolin	38966-21-1	+	20	20	0.06	+	20	20	0.06			-	+	-			
AraC	147-94-4	-		1400	5.76	-		1400	5.76		-	-	+		+	+	+
5-Azacytidine	320-67-2	++	625	1250	5.12	++	625	1250	5.12		+/-	+	+		+	+	
AZT	30516-87-1	+	3675	7350	27.5	-		7350	27.5		+/-	-	+			+	+
Benzaldehyde	100-52-7	++	2625	5220	49.2	+	1313	5220	49.2		+/-	-	+				+
Benzo(a)pyrene	50-32-8	-		26	0.10	+		26	0.10		+	+/- (MA)	+ (MA)	+ (MA)	+ (MA)	+/-	+
Benzoyl Chloride	98-88-4	++	1515	12120	86.22	+	304	12120	86.22			+					+
Bleomycin Sulfate	9041-93-4	++	5	5	0.0035	++	2.5	5	0.0035		+	+/-	+	+	+	+	+
Cadmium Chloride	10108-64-2	+	2	2	0.011	-		2	0.011		+/-	-		-	+/-	+/-	+
Caffeine	58-08-2	++	266	531	2.73	-		531	2.73		-	-	+	-		+/-	+
Catechol	120-80-9	++	177	660	7.99	++	599	660	7.99		+/-	-	+	-		-	
Cefotaxime	64485-93-4	++	238	238	0.50	++	238	238	0.50								
Chlorambucil	305-03-3	+	125	250	0.82	-		250	0.82		+	+/-				+	+
Chloramphenicol	56-75-7	+	162	162	0.50	++	10.1	162.0	0.50			-	+	-		-	
Chromomycin A3	7059-24-7	-		10	0.0085	-		10.0	0.0085			-		-			
Cimetidine	51481-61-9	+	125	125	0.50	++	31.3	125	0.50		-						
Cisplatin (without DMSO)	15663-27-1	++	2.34	150	0.50	+	18.75	37.5	0.12		+	+/-		+			
Colchicine	64-86-8	-		853	2.14	+	427	853	2.14		+	-	-		+	+/-	
Crotonaldehyde	4170-30-3	++	26.5	1692	24.1	++	85.2	169	2.41		+	+			+		
Cumene Hydroperoxide	80-15-9	++	65.4	519	3.41	-		104	0.68			+		+			
Cycloheximide	66-81-9	++	0.03	1	0.0036	-		1450	5.15			-	+			+	
Daunorubicin	23541-50-6	++	0.5	250	0.44	++	0.25	0.5	0.0009		+	+/-		+			
3,5-Dichlorophenol	591-35-5	++	12.5	25	0.15	-		25	0.1534			-					
Dicumyl Peroxide	80-43-3	++	0.3	0.3	0.0011	-		0.3	0.0011			-					
Dieldrin	60-57-1	-	19.8	39.6	0.10	-		39.6	0.10		+/-	-	+	-		+	+
Diethylamino-4-methylcoumarin	91-44-1	++	59	118	0.51	-		118	0.51			-					
1,2-Dimethylhydrazine HCl	306-37-6	++	330	1330	10.00	++	665	1330	10.00		+	-				+	
Econazole Nitrate	24169-02-6	++	25	50	0.11	++	2.5	5	0.011			-			+/-		
Ellipticine	519-23-3	++	4.39	562	2.28	++	12.5	562	2.280			+/-	+/-				
1,2-Epoxybutane	106-88-7	+	8370	16740	232	+	16740	16740	232		+/-	+/-	+	+		-	
Ethidium Bromide	1239-45-8	-		12.5	0.032	++	6.25	6.25	0.016			+/- (MA)	+/-	+ (MA)	+	+	+
7-Ethoxycoumarin	31005-02-4	-		129	0.68	++	32.3	129	0.68								
Ethyl Acrylate	140-88-5	++	4605	18420	184.0	-		18420	183.98		+/-	-	+				+
Ethylenediamine	107-15-3	-	224.8	450	7.49	-		450	7.49		-	-					

Fig.20 [2/3]

Ethyl methanesulfonate	62-50-0	-		100	0.81	++	25	100	0.81		+	+/-	+			+		
Etoposide	33419-42-0	-		125	0.21	+	125	125	0.21			-	+			+		
Furazolidone	67-45-8	++	14.1	113.0	0.50	++	28.2	113	0.50			+		+				
Hexamethylenetetramine	100-97-0	++	150.0	300.0	2.14	+	150	300	2.14			+/-						
Hydrazine monohydrate	7803-57-8	++	3.2	204	4.08	++	4	4	0.08		+	+/-	+	+ (MA)				
Hydrogen Peroxide	7722-84-1	++	4.29	550	16.17	+	68.5	137.5	4.04		+	+/-	+	+				+
Hydroquinone	123-31-9	+	55.0	55	0.50	-		55	0.50		+/-	+/-	+		+	+		
8-Hydroxyquinoline	148-24-3	++	0.13	0.5	0.0034	-		0.125	0.0009		-	+/- (MA)	+	-		-		
Hydroxyurea	127-07-1	++	1187	19015	250.0	++	593	19015	250.00			+/-	+	+		-	+	
ICR191 Acridine Mutagen	17070-45-0	++	1.56	12.5	0.028	-		12.5	0.028	FP		+	+					
Isobutyl Nitrite	542-56-3	++	127	508	3.90	-		508	3.90		+	+/- (MA)	+			+		
Mechlorethamine HCl	55-86-7	-		1920	12	+	96	96	0.62		+	+/-				+		
Methapyrilene HCl	135-23-9	++	9.0	144	0.48	++	71	144	0.48	FP	+/-	-	+ (MA)	+		+		
Methyl Carbamate	598-55-0	-		38	0.50	-		37.5	0.50		+/-	-	+			-		
Methyl Methacrylate	80-62-6	-		94300	941.9	-		94300	941.67	FP	-	-	+					
Methyl methanesulfonate	66-27-3	+	16.25	32.5	0.30	++	1.02	33	0.30		+	+/-	+	+	+	+	+	
Methyl viologen	1910-42-5	+	100	142	0.55	+	40	40	0.16		+	-	+	+				
Mitomycin C	50-07-7	+	100	200	0.60	+	200	200	0.60		+	+/-	+	+	+	+		
MNNG	70-25-7	++	0.57	74	0.50	++	0.125	0.500	0.0034		+	+	+	+	+	-	+	
Nalidixic Acid	389-08-2	++	58	116	0.60	++	0.25	11.5	0.05		+/-	+/-	-	+	-			
1-Naphthylamine	134-32-7	++	33.3	133	0.93	+	10	10	0.070		-	+/- (MA)		+	-	+/-	+	
Netropsin	18133-22-7	++	15.6	250	0.50	++	7.8	250	0.50									
Neutral Red	553-24-2	++	0.18	2.88	0.010	++	0.72	2.88	0.010			+ (MA)						+
Nicotine	54-11-5	++	5050	5050	31.1	-		5050	31.1		-	-		-				-
Nitrobenzene	98-95-3	-		12040	97.8	-		12040	97.8		+	-						-
Nitrofurantoin	67-20-9	+	115	115	0.48	-		115	0.48		+/-	+/-	+	+		-		+
4-Nitroquinoline-N-oxide	56-57-5	++	0.13	1.0	0.005	++	0.13	1.0	0.005		+	+	+					
N-Nitrosodimethylamine	62-75-9	++	6288	50300	679	-		50300	679		+/-	+/-	+ (MA)	+ (MA)	+	+/-		
N-Nitrosodiphenylamine	86-30-6	-		99	0.50	-		99	0.50		+/-	+/- (MA)			-	-	-	
N-Nitroso-N-ethyl urea	759-73-9	++	1450	2900	25	++	127	203	1.73		+	+				+		
N-Nitroso-N-methyl urea	684-93-5	-		104	1.01	++	52	104	1.01		+	+			+	+		
Noscapine	128-62-1	-		2250	5.00	-		2250	5.00									+ (p)
4,4-Oxydianiline	101-80-4	-	82.5	165	0.82	-		165	0.82		+	+/-	+			+		
Phenol	108-95-2	-	235	470	4.99	-		470	4.99		-	-		-	+/-	+/-		
Phenylmethylsulfonylfluoride	329-98-6	-		87	0.50	-		87	0.50									
Phleomycin	11006-33-0	++	3.13	12.5		++	12.5	12.5										
Phthalic acid, bis(2-ethylhexyl) ester	117-81-7	+	98100	98100	251	++	4905	98100	251		+	-		-	+/-	-		
Proflavin Hemisulfate	553-30-0	++	4.0	127.0	0.49	++	1	2	0.008	FP	+	+/- (MA)	+					
Psoralen	66-97-7	-		93.1	0.50	+	93.1	93.1	0.500			-		+	+			
Quercetin	117-39-5	-	108	108.0	0.32	-		108	0.32		+/-	+/-	+ (MA)	+ (MA)	+	-	+	
Safrole	94-59-7	++	273	545	3.36	++	34	545	3.36		+	-	+ (MA)	-	-	+/-	+	
Sodium Azide	26628-22-8	++	4.06	32.5	0.50	++	0.66	5.3	0.082		-	+/-	+	-			+	
Sodium Selenite	10102-18-8	-	63	1000	5.78	-		1000	5.78			+			-		-	
Streptonigrin	3930-19-6	++	0.39	25	0.049	++	0.39	25	0.049			+		+			+	
Sulfamethoxazole	723-46-6	++	5	20	0.08	++	2.5	20	0.079		+	-						
Sulfaphenazole	526-08-9	++	19.4	155	0.49	-		5	0.016									
Sulfisoxazole	127-69-5	++	16.9	135	0.51	-		10	0.037		-	-	+			-		
Taxol	33069-62-4	+	25	25	0.03	-		25	0.029			-			+	+		
Tetracycline HCl	64-75-5	-	267	267	0.60	-		267	0.60		-	-	-					+

Thiourea	62-56-6	++	10050	20100	264	++	1256	5000	65.7		+	-		-	+/-		
Titanium Dioxide	13463-67-7	-		179	2.24	-		179	2.24		+/-	-	-		-	+	
Trichloroacetonitrile	545-06-2	++	3.6	28.8	0.20	++	1.8	28.8	0.20			+/-				-	
Trilolyl Phosphate	1330-78-5	-		572	1.55	++	143	572	1.55		-	-	-				
Urethane	51-79-6	++	10300	20600	231	++	10300	10300	115.60		+	-		-	+	+	+
Vanillin	121-33-5	++	21.25	170	1.12	-		152	1.00		-	-			-	-	-
Vinblastine	143-67-9	++	62.5	250	0.28	-		250	0.28		-	-				+	+

KEY	
-	Negative
+	Positive
++	Strong Positive
+/-	Results vary between reports
MA	Metabolic activation required to obtain positive result.
LEC	Lowest effective concentration
FP	Fluorescence polarisation used to reveal the result.
(p)	Polyploid problems increase significantly

FIG.21

TEST COMPOUND	GSA	Ames	S9	TEST COMPOUND	GSA	Ames	S9
2-Amino-4-nitrophenol	++	+	+	2-Acetamidofluorene	-	+	+
Ethidium Bromide	++	+	+	2-Aminoanthracene	-	+	+
Neutral Red	++	+	+	o-Anisidine	-	+	+
Proflavin Hemisulfate	++	+	+	8-Hydroxyquinoline	-	+	+
5-Azacytidine	++	+	-	Isobutyl Nitrite	-	+	+
Bleomycin Sulfate	++	+	-	N-Nitrosodiphenylamine	-	+	+
Crotonaldehyde	++	+	-	9-Aminoacridine	-	+	-
Daunorubicin	++	+	-	Chlorambucil	-	+	-
Ellipticine	++	+	-	Cumene Hydroperoxide	-	+	-
Ethyl methanesulfonate	++	+	-	Hydroquinone	-	+	-
Furazolidone	++	+	-	ICR191 Acridine Mutagen	-	+	-
Hydrazine monohydrate	++	+	-	Nitrofurantoin	-	+	-
Hydroxyurea	++	+	-	N-Nitrosodimethylamine	-	+	-
Methyl methanesulfonate	++	+	-	4,4-Oxydianiline	-	+	-
MNNG	++	+	-	Quercetin	-	+	-
Nalidixic Acid	++	+	-	Sodium Selenite	-	+	-
4-Nitroquinoline-N-oxide	++	+	-	Acetylsalicylic Acid	-	-	
N-Nitroso-N-ethyl urea	++	+	-	Actinomycin D	-	-	
N-Nitroso-N-methyl urea	++	+	-	4-Aminophenol	-	-	
Sodium Azide	++	+	-	Ampicillin (Na salt)	-	-	
Streptonigrin	++	+	-	Aniline	-	-	
Trichloroacetonitrile	++	+	-	AraC	-	-	
Benzo(a)pyrene	+	+	+	AZT	-	-	
1-Naphthylamine	+	+	+	Cadmium Chloride	-	-	
Benzoyl Chloride	+	+	-	Caffeine	-	-	
Cisplatin (without DMSO)	+	+	-	Chromomycin A3	-	-	
1,2-Epoxybutane	+	+	-	Cycloheximide	-	-	
Hexamethylenetetramine	+	+	-	3,5-Dichlorophenol	-	-	
Hydrogen Peroxide	+	+	-	Dicumyl Peroxide	-	-	
Mechlorethamine HCl	+	+	-	Dieldrin	-	-	
Mitomycin C	+	+	-	Diethylamino-4-methylcoumarin	-	-	
3-Amino-1,2,4-triazole	+	-		Ethyl Acrylate	-	-	
Aphidicolin	+	-		Ethylenediamine	-	-	
Benzaldehyde	+	-		Methyl Carbamate	-	-	
Colchicine	+	-		Methyl Methacrylate	-	-	
Etoposide	+	-		Nicotine	-	-	
Methyl viologen	+	-		Nitrobenzene	-	-	
Psoralen	+	-		Phenol	-	-	
Catechol	++	-		Sulfisoxazole	-	-	
Chloramphenicol	++	-		Taxol	-	-	
1,2-Dimethylhydrazine HCl	++	-		Tetracycline HCl	-	-	
Econazole Nitrate	++	-		Titanium Dioxide	-	-	
Methapyrilene HCl	++	-		Vanillin	-	-	
Phthalic acid bis(2-ethylhexyl) ester	++	-		Vinblastine	-	-	
Safrole	++	-					
Sulfamethoxazole	++	-		S9 column records requirement for Ames result:			
Thiourea	++	-					
Tritolyl Phosphate	++	-		+ S9 required			
Urethane	++	-		- S9 not required			

Table 3

FIG. 22

Greenrack loading sequence

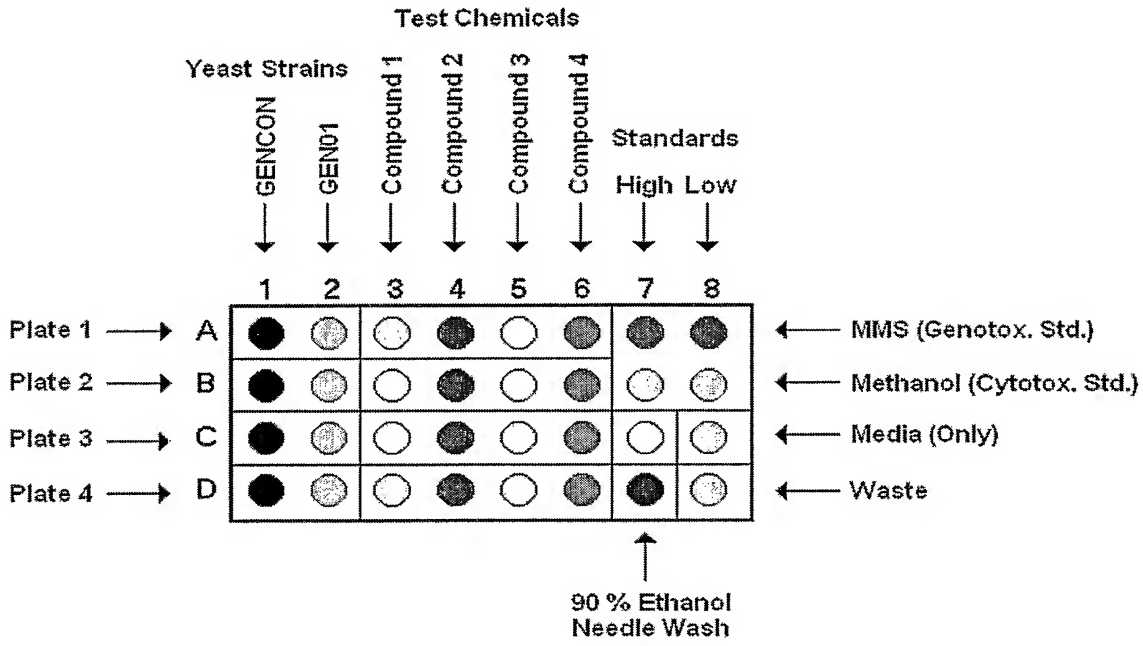


FIG. 23

Microplate layout

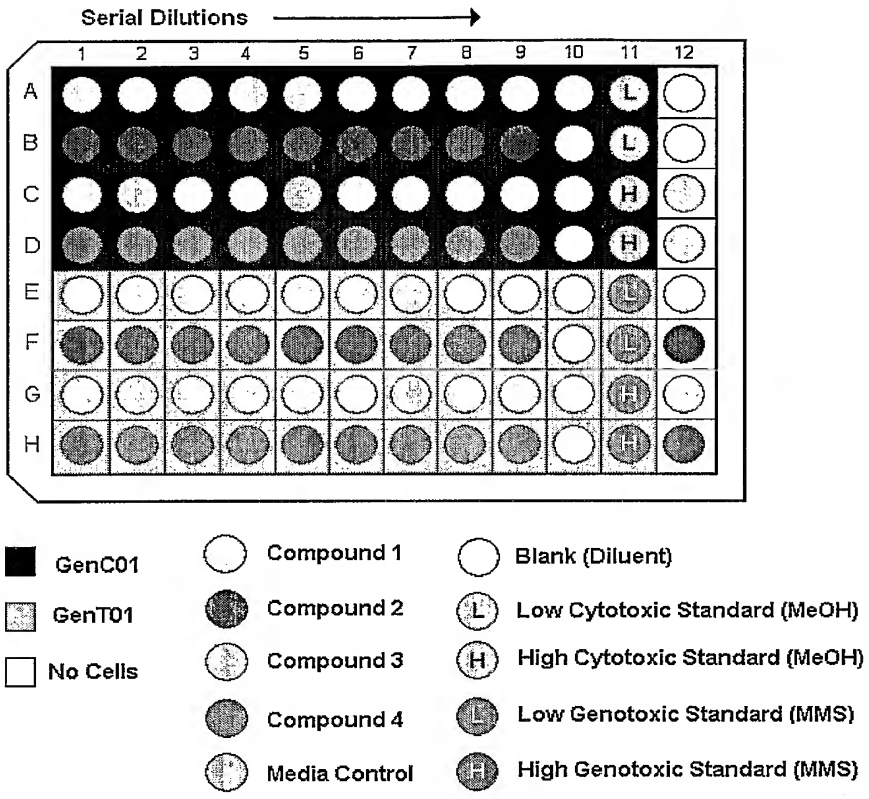


FIG. 24

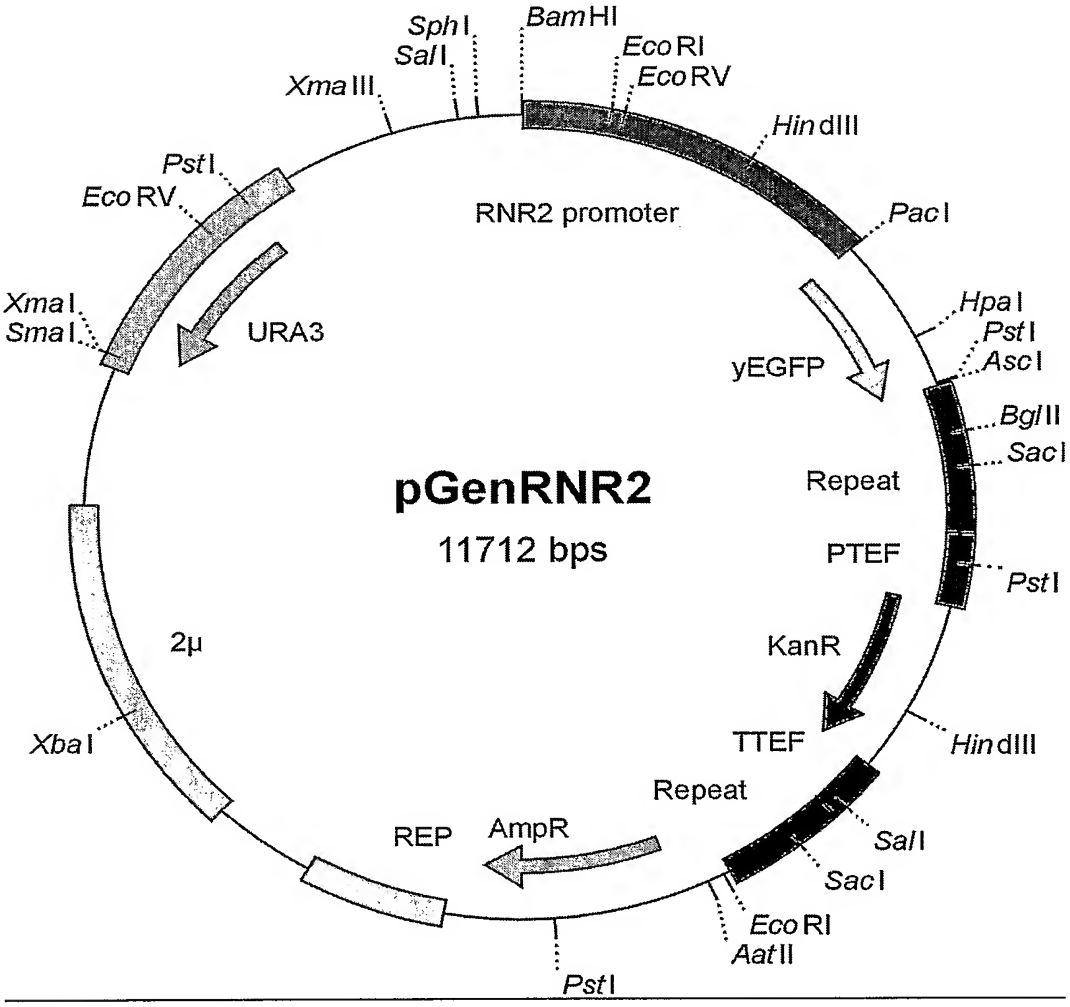

pGenRNR2
11712 bps
RNR2 promoter
yEGFP
Repeat
PTEF
KanR
TTEF
Repeat
AmpR
REP
2µ
URA3
Sph I
Sal I
Xma III
Bam HI
Eco RI
Eco RV
Hin dIII
Pac I
Hpa I
Pst I
Asc I
Bgl II
Sac I
Pst I
Hin dIII
Sal I
Sac I
Eco RI
Aat II
Pst I
Xba I
Xma I
Sma I
Eco RV
Pst I

FIG. 25

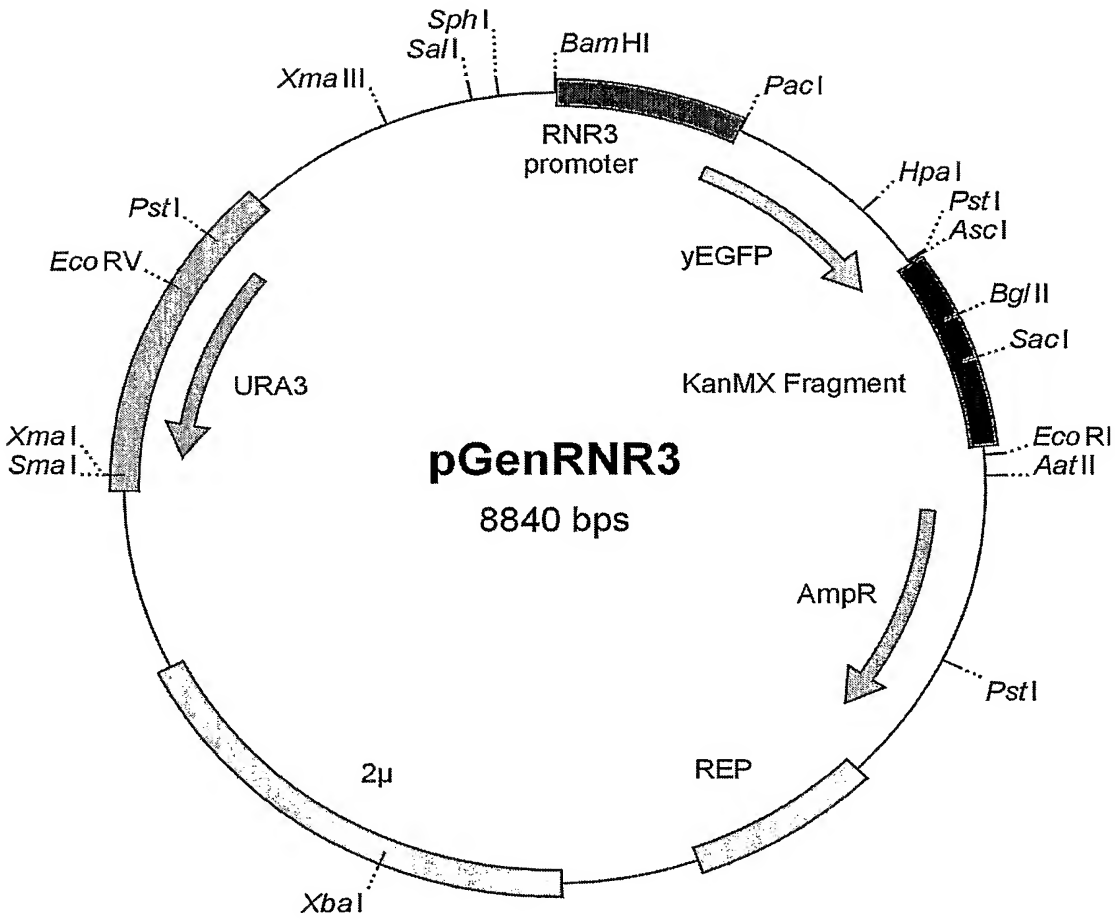

Sph I
Sal I
Bam HI
Xma III
Pac I
RNR3
promoter
Hpa I
Pst I
Asc I
Pst I
Eco RV
yEGFP
Bgl II
Sac I
URA3
KanMX Fragment
Xma I
Sma I
pGenRNR3
8840 bps
Eco RI
Aat II
AmpR
Pst I
2µ
REP
Xba I

FIG. 26

pGenRNR2

GGATCCGTACCTTCCAGCATTGTCCTTCTGAGAAAACAAAAATGGAAGATGTTGTGAAAATGCAGTAAGTGA
CAATAAGCTAGCGACTTTTCAATGATCGCAGATTAAGTATATATACACTTTAGGTCTCCTTCTCTTCTTTAC
TAGCTACGAAAAAACAAAGAAAAAAAAAACAAAATAGAGGGAAAAAAGATATACGTATAAAAATATGCTCCA
GGGGAGGTTCGAACTCTCGACCTTCAGATTATGAGACTGACGCTCTTCCTACTGAGCTACTGAAGCACTATT
TCTTGTTCCAACCGAGGAATAATACTAGACCTGTAGAAAAGTACCGCTTATCTAATATCCAACTTTTTCTTC
TGTTCCCTTTAGTGAATTCAAGATGAAGGAACATAACAGAACGGAAAACAAAAAAAAAAGGGAAGAAAAAAGA
TATCACCACCAATTGGCACGCTCGCAACGGTTTCGGGCCAGATTCATGGCCCTGTGGGGTAGAAGTGCCTAC
AGACAAAAAACCGTTACGTCCCGCCTCACCAAGCATCGACACACAAGCCCTCTCAACCTCCTTAATTTCTTT
ATCTTTTTTTCCCTTCACCTCTGTGCTCTTTTTTTAGCCACAGCTGTGCATTACGCTTGCTTTGCTTGTTTT
GTTCTCTTTCGCCATGCCATATTACTACTTTTCTACTTACTGGCACGAGGCCGCGCAAGCCAGATCCAAGCA
CGCCAGAAAGTGCGGTGTACTGGTGCAACCTTTTATCGATTTCTCATCATATCGCTATCGCAACGGGCATGG
GCGACGAAAAAGCCAATCTAAAGCGTTGTCTTCCTCGTTCCTCGCGCACGCGATTTTGCGTTCTGCATAGGA
AGCCGAAGTCGAACAAGAAGCAGGCAAAGTTTAGAGCACTGCCCCTCCGCACTCAAAAAAGAAAAAACTAGG
AGGAAAATAAAATTCTCAACCACACAAACACATAAACACATACAAATACAAATACAAGCTTATTTACTTGAC
ATCGCGCGATCTTCCACTATTCAGCGCCGTCCGCCCTCTCTCGTGTTTTTTGTTTACGCGACAACTATGCGA
AATCCGGAGCAACGGGCAACCGTTTGGGGAAAGACCACACCCACGCGCGATCGCCATGGCAACGAGGTCGCA
CACGCCCCACACCCAGACCTCCCTGCGAGCGGGCATGGGTACAATGTCCCCGTTGCCACAGAGACCACTTCG
TAGCACAGCGCAGAGCGTAGCGTGTTGTTGCTGCTGACAAAAGAAAATTTTTCTTAGCAAAGCAAAGGAGGG
GAAGCACGGGCAGATAGCACCGTACCATACCCTTGGAAACTCGAAATGAACGAAGCAGGAAATGAGAGAATG
AGAGTTTTGTAGGTATATATAGCGGTAGTGTTTGCGCGTTACCATCATCTTCTGGATCTATCTATTGTTCTT
TTCCTCATCACTTTCCCCTTTTTCGCTCTTCTTCTTGTCTTTTATTTCTTTCTTTTTTTTAATTGTTCCCTC
GATTGGCTATCTACCAAAGAATCCAAACTTAATACACGTATTTATTTGTCCAATTACCATGTTAATTAACTC
TAAAGGTGAAGAATTATTCACTGGTGTTGTCCCAATTTTGGTTGAATTAGATGGTGATGTTAATGGTCACAA
ATTTTCTGTCTCCGGTGAAGGTGAAGGTGATGCTACTTACGGTAAATTGACCTTAAAATTTATTTGTACTAC
TGGTAAATTGCCAGTTCCATGGCCAACCTTAGTCACTACTTTCGGTTATGGTGTTCAATGTTTTGCGAGATA
CCCAGATCATATGAAACAACATGACTTTTTCAAGTCTGCCATGCCAGAAGGTTATGTTCAAGAAAGAACTAT
TTTTTTCAAAGATGACGGTAACTACAAGACCAGAGCTGAAGTCAAGTTTGAAGGTGATACCTTAGTTAATAG
AATCGAATTAAAAGGTATTGATTTTAAAGAAGATGGTAACATTTTAGGTCACAAATTGGAATACAACTATAA
CTCTCACAATGTTTACATCATGGCTGACAAACAAAAGAATGGTATCAAAGTTAACTTCAAAATTAGACACAA
CATTGAAGATGGTTCTGTTCAATTAGCTGACCATTATCAACAAAATACTCCAATTGGTGATGGTCCAGTCTT
GTTACCAGACAACCATTACTTATCCACTCAATCTGCCTTATCCAAAGATCCAAACGAAAAGAGAGACCACAT
GGTCTTGTTAGAATTTGTTACTGCTGCTGGTATTACCCATGGTATGGATGAATTGTACAAATAACTGCAGGG
CGCGCCACTTCTAAATAAGCGAATTTCTTATGATTTATGATTTTTATTATTAAATAAGTTATAAAAAAAATA
AGTGTATACAAATTTTAAAGTGACTCTTAGGTTTTAAAACGAAAATTCTTATTCTTGAGTAACTCTTTCCTG
TAGGTCAGGTTGCTTTCTCAGGTATAGTATGAGGTCGCTCTTATTGACCACACCTCTACCGGCAGATCCGCT
AGGGATAACAGGGTAATATAGATCTGCCCGCCGGGAAGGCGAACCCGATCGGATGCATCCTCTCTGCTGCCA
TGATGCTGAAGTTGTCGTTGAACATGGTTGCTGCCGGCGAGGCGGTCGAGCAGGCAGTGCAGGAGGTGTTGG
ACTCGGGAGTCAGAACGGGCGACCTGCTCGGCTCGAGCTCCACTTCGGAGGTTGGCGACGCCATTGCGCTTG
CAGTTAAGGAAGCCTTGCGCAGGCAATCCGCAGCTGGTCTGAGCTAGCCTCGAGGACCCTTCTCTTTAGACT
ATTCTACTCTTATGCACGTAAAAAATTCTAGGAAATATGTATTAACTAGGAGTAAAATAACCGGCTAGTGGC
ATTCATATAGCCGTCTGTTTACATCTACATCACACATTTCGAGTGTATATCTCGCAACGTTGGCGTTAAATA
GGCAGTCAATGGCCCGACCATTCTATGGTGTTTAGGTCGATGCCATCTTTGTACAGCTTGCCTCGTCCCCGC
CGGGTCACCCGGCCAGCGACATGGAGGCCCAGAATACCCTCCTTGACAGTCTTGACGTGCGCAGCTCAGGGG
CATGATGTGACTGTCGCCCGTACATTTAGAGCAAAAATTACGGCTCCTCGCTGCAGACCTGCGAGCAGGGAA
ACGCTCCCCTCACAGACGCGTTGAATTGTCCCCACGCCGCGCCCCTGTAGAGAAATATAAAAGGTTAGGATT
TGCCACTGAGGTTCTTCTTTCATATACTTCCTTTTAAAATCTTGCTAGGATACAGTTCTCACATCACATCCG
AACATAAACAACCATGGGTAAGGAAAAGACTCACGTTTCGAGGCCGCGATTAAATTCCAACATGGATGCTGA
TTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAATCTATCGATTGTATGGGAA
GCCCGATGCGCCAGAGTTGTTTC

TGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAAT
TTATGCCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCC
CCGGCAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAG
TGTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCGATCGCGTATTTCGTCTCG
CTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGGC
CTGTTGAACAAGTCTGGAAAGAAATGCATAAGCTTTTGCCATTCTCACCGGATTCAGTCGTCACTCATGGTG
ATTTCTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAA
TCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAAC
GGCTTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGT
TTTTCTAATCAGTACTGACAATAAAAAGATTCTTGTTTTCAAGAACTTGTCATTTGTATAGTTTTTTTATAT
TGTAGTTGTTCTATTTTAATCAAATGTTAGCGTGATTTATATTTTTTTTCGCCTCGACATCATCTGCCCAGA
TGCGAAGTTAAGTGCGCAGAAAGTAATATCATGCGTCAATCGTATGTGAATGCTGGTCGCTATACTGCTGTC
GATTCGATACTAACGCCGCCATCCAGTGTCGACTAGGGTTGCTGCCATCGGCCTCGCTCGCGTCTTTGCCGG
ATAGCAAGAGCGCCTTTGGCCTCTACGAGCCCTGCCACGGCTCTGCGCCCGATCTGCCCGCCGGGAAGGCGA
ACCCGATCGGATGCATCCTCTCTGCTGCCATGATGCTGAAGTTGTCGTTGAACATGGTTGCTGCCGGCGAGG
CGGTCGAGCAGGCAGTGCAGGAGGTGTTGGACTCGGGAGTCAGAACGGGCGACCTGCTCGGCTCGAGCTCCA
CTTCGGAGGTTGGCGACGCCATTGCGCTTGCAGTTAAGGAAGCCTTGCGCAGGCAATCCGCAGCTGGTCTGA
GCTAGCCTCGAGGACCCTTCTCTTTAGACTATTCTACTCTTATGCACGTAAAAAATTCTAGGAAATATGTAT
TAACTAGGAGTAAAATAACCGGCTAGTGGCATTCATATAGCCGTCTGTTTACATCTACATCACACATTTCGA
GTGTATATCTCGCAACGTTGGCGTTAAATAGGCAGTCAATGGCCCGACCATTCTATGGTGTTTAGGTCGATG
CCATCTTTGTACGTTTAGCTTATCGATGATAAGCTGTCAAACATGAGAATTCTTGAAGACGAAAGGGCCTCG
TGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGG
GAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAAT
AACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTA
TTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTG
AAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTC
GCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTGTTG
ACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCA
CAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACA
CTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGG
ATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCA
CGATGCCTGCAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGC
AACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCT
GGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATG
GTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGA
TCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGA
TTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAA
TCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATC
CTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGG
ATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTC
TAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCC
TGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGG
ATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCG
AACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATC
CGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATA
GTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTAT
GGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTC
CTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCC
GAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGC
ATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCC
AGTATACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGC
GCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTG
TCAGAGGTTTTCACCGTCA

TCACCGAAACGCGCGAGGCAGAGCTTTGAAGAAAAATGCGCCTTATTCAATCTTTGCTATAAAAAATGGCCC
AAAATCTCACATTGGAAGACATTTGATGACCTCATTTCTTTCAATGAAGGGCCTAACGGAGTTGACTAATGT
TGTGGGAAATTGGAGCGATAAGCGTGCTTCTGCCGTGGCCAGGACAACGTATACTCATCAGATAACAGCAAT
ACCTGATCACTACTTCGCACTAGTTTCTCGGTACTATGCATATGATCCAATATCAAAGGAAATGATAGCATT
GAAGGATGAGACTAATCCAATTGAGGAGTGGCAGCATATAGAACAGCTAAAGGGTAGTGCTGAAGGAAGCAT
ACGATACCCCGCATGGAATGGGATAATATCACAGGAGGTACTAGACTACCTTTCATCCTACATAAATAGACG
CATATAAGTACGCATTTAAGCATAAACACGCACTATGCCGTTCTTCTCATGTATATATATATACAGGCAACA
CGCAGATATAGGTGCGACGTGAACAGTGAGCTGTATGTGCGCAGCTCGCGTTGCATTTTCGGAAGCGCTCGT
TTTCGGAAACGCTTTGAAGTTCCTATTCCGAAGTTCCTATTCTCTAGAAAGTATAGGAACTTCAGAGCGCTT
TTGAAAACCAAAAGCGCTCTGAAGACGCACTTTCAAAAAACCAAAAACGCACCGGACTGTAACGAGCTACTA
AAATATTGCGAATACCGCTTCCACAAACATTGCTCAAAAGTATCTCTTTGCTATATATCTCTGTGCTATATC
CCTATATAACCTACCCATCCACCTTTCGCTCCTTGAACTTGCATCTAAACTCGACCTCTACATTTTTTATGT
TTATCTCTAGTATTACTCTTTAGACAAAAAAATTGTAGTAAGAACTATTCATAGAGTGAATCGAAAACAATA
CGAAAATGTAAACATTTCCTATACGTAGTATATAGAGACAAAATAGAAGAAACCGTTCATAATTTTCTGACC
AATGAAGAATCATCAACGCTATCACTTTCTGTTCACAAAGTATGCGCAATCCACATCGGTATAGAATATAAT
CGGGGATGCCTTTATCTTGAAAAAATGCACCCGCAGCTTCGCTAGTAATCAGTAAACGCGGGAAGTGGAGTC
AGGCTTTTTTTATGGAAGAGAAAATAGACACCAAAGTAGCCTTCTTCTAACCTTAACGGACCTACAGTGCAA
AAAGTTATCAAGAGACTGCATTATAGAGCGCACAAAGGAGAAAAAAAGTAATCTAAGATGCTTTGTTAGAAA
AATAGCGCTCTCGGGATGCATTTTTGTAGAACAAAAAAGAAGTATAGATTCTTTGTTGGTAAAATAGCGCTC
TCGCGTTGCATTTCTGTTCTGTAAAAATGCAGCTCAGATTCTTTGTTTGAAAAATTAGCGCTCTCGCGTTGC
ATTTTTGTTTTACAAAAATGAAGCACAGATTCTTCGTTGGTAAAATAGCGCTTTCGCGTTGCATTTCTGTTC
TGTAAAAATGCAGCTCAGATTCTTTGTTTGAAAAATTAGCGCTCTCGCGTTGCATTTTTGTTCTACAAAATG
AAGCACAGATGCTTCGTTCTGCGGTAAAGCTCATCAGCGTGGTCGTGAAGCGATTCACAGATGTCTGCCTGT
TCATCCGCGTCCAGCTCGTTGAGTTTCTCCAGAAGCGTTAATGTCTGGCTTCTGATAAAGCGGGCCATGTTA
AGGGCGGTTTTTTCCTGTTTGGTCACTGATGCCTCCGTGTAAGGGGGATTTCTGTTCATGGGGGTAATGATA
CCGATGAAACGAGAGAGGATGCTCACGATACGGGTTACTGATGATGAACATGCCCGGTTACTGGAACGTTGT
GAGGGTAAACAACTGGCGGTATGGATGCGGCGGGACCAGAGAAAAATCACTCAGGGTCAATGCCAGCGCTTC
GTTAATACAGATGTAGGTGTTCCACAGGGTAGCCAGCAGCATCCTGCGATGCAGATCCGGAACATAATGGTG
CAGGGCGCTGACTTCCGCGTTTCCAGACTTTACGAAACACGGAAACCGAAGACCATTCATGTTGTTGCTCAG
GTCGCAGACGTTTTGCAGCAGCAGTCGCTTCACGTTCGCTCGCGTATCGGTGATTCATTCTGCTAACCAGTA
AGGCAACCCCGCCAGCCTAGCCGGGTCCTCAACGACAGGAGCACGATCATGCGCACCCGTGGCCAGGACCCA
ACGCTGCGGGGGGGGGGGGGGTTTTCTTTCCAATTTTTTTTTTTTCGTCATTATAGAAATCATTACGACCGA
GATTCCCGGGTAATAACTGATATAATTAAATTGAAGCTCTAATTTGTGAGTTTAGTATACATGCATTTACTT
ATAATACAGTTTTTTAGTTTTGCTGGCCGCATCTTCTCAAATATGCTTCCCAGCCTGCTTTTCTGTAACGTT
CACCCTCTACCTTAGCATCCCTTCCCTTTGCAAATAGTCCTCTTCCAACAATAATAATGTCAGATCCTGTAG
AGACCACATCATCCACGGTTCTATACTGTTGACCCAATGCGTCTCCCTTGTCATCTAAACCCACACCGGGTG
TCATAATCAACCAATCGTAACCTTCATCTCTTCCACCCATGTCTCTTTGAGCAATAAAGCCGATAACAAAAT
CTTTGTCGCTCTTCGCAATGTCAACAGTACCCTTAGTATATTCTCCAGTAGCTAGGGAGCCCTTGCATGACA
ATTCTGCTAACATCAAAAGGCCTCTAGGTTCCTTTGTTACTTCTTCCGCCGCCTGCTTCAAACCGCTAACAA
TACCTGGGCCCACCACACCGTGTGCATTCGTAATGTCTGCCCATTCTGCTATTCTGTATACACCCGCAGAGT
ACTGCAATTTGACTGTATTACCAATGTCAGCAAATTTTCTGTCTTCGAAGAGTAAAAAATTGTACTTGGCGG
ATAATGCCTTTAGCGGCTTAACTGTGCCCTCCATGGAAAAATCAGTCAAGATATCCACATGTGTTTTTAGTA
AACAAATTTTGGGACCTAATGCTTCAACTAACTCCAGTAATTCCTTGGTGGTACGAACATCCAATGAAGCAC
ACAAGTTTGTTTGCTTTTCGTGCATGATATTAAATAGCTTGGCAGCAACAGGACTAGGATGAGTAGCAGCAC
GTTCCTTATATGTAGCTTTCGACATGATTTATCTTCGTTTCCTGCAGGTTTTTGTTCTGTGCAGTTGGGTTA
AGAATACTGGGCAATTTCATGTTTCTTCAACACCACATATGCGTATATATACCAATCTAAGTCTGTGCTCCT
TCCTTCGTTCTTCCTTCTGCTCGGAGATTACCGAATCAAAAAAATTTCAAAGAAACCGGAATCAAAAAAAAG
AACAAAAAAAAAAAAAGATGAATTGAAACCCCCCCCCCCCCCGATGCGCCGCGTGCGGCTGCTGGAGATGGCG
GACGCGATGGATATGTTCTGCCAAGGGTTGGTTTGCGCATTCACAGTTCTCCGCAAGAATTGATTGGCTCCA
ATTCTTGGAGTGGTGAATCCGTTAGCGAGGTGCCGCCGGCTTCCATTCAGGTCGAGGTGGCCCGGCTCCATG
CACCGCGACGCAACGCGGGGAGGCAGACAAGGTATAGGGCGGCGCCTACAATCCATGCCAACCCGTTCCATG
TGCTCGCCGAGGCGGCATAAATCGCCGTGACGATCAGCGGTCCAGTGATCGAAGTTAGGCTGGTAAGAGCCG
CGAGCGATCCTTGAAGCTGTCCCTGATGGTCGTCATCTACCTGCCTGGACAGCATGGCCTGCAACGCGGGCA
TCCCGATGCCGCCGGAAG

CGAGAAGAATCATAATGGGGAAGGCCATCCAGCCTCGCGTCGCGAACGCCAGCAAGACGTAGCCCAGCGCGT
CGGCCGCCATGCCGGCGATAATGGCCTGCTTCTCGCCGAAACGTTTGGTGGCGGGACCAGTGACGAAGGCTT
GAGCGAGGGCGTGCAAGATTCCGAATACCGCAAGCGACAGGCCGATCATCGTCGCGCTCCAGCGAAAGCGGT
CCTCGCCGAAAATGACCCAGAGCGCTGCCGGCACCTGTCCTACGAGTTGCATGATAAAGAAGACAGTCATAA
GTGCGGCGACGATAGTCATGCCCCGCGCCCACCGGAAGGAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCG
GTCGACGCTCTCCCTTATGCGACTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCG
CCGCCGCAAGGAATGGTGCATGCAAGGAGATGGCGCCCAACAGTCCCCCGGCCACGGGGCCTGCCACCATAC
CCACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCCCCATCGGTGATGTCGGCGATAT
AGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATGCGTCCGGCGTAGA

FIG. 27

PGenRNR3

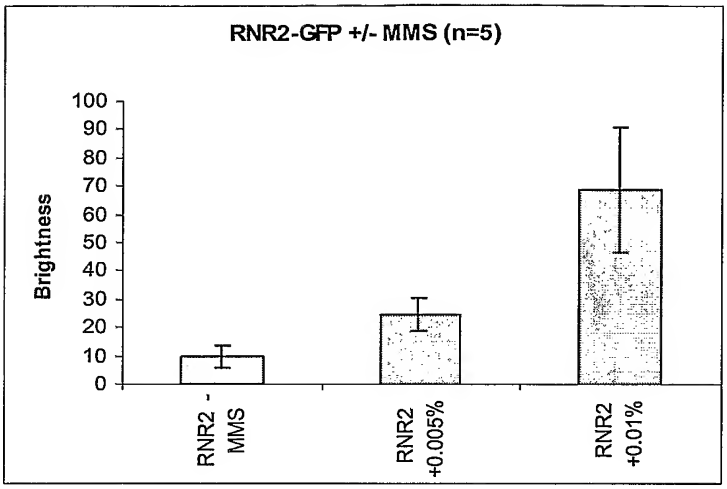
GGATCCAGAAACAAGAGAAGGTAACAAGCACATAAAAAATCAGCACATACGTACATACATAAGAATGAATCG
CACGCACGCGTAAACATTTATCATTTAATCTTCAGTTGTTAGATAAAAAAAAAAAGAAAAGAAAAGAAAGTG
AAGGCTTGTTTCAGTTTGAACTAGGTAGCAGAGCAAGCCCTCGTTCTTGGCTGCTAATTTTCCTAAAGTAGT
AAAAAAAGCCAAGTTATCTGCCTACGGTTGTCACAGCAACATTGCGTGCCGTTGTTCTTTTGTTTTTTTTTT
TTTTTTTTTTTCGTGGTTGTCGCAGCAACGACACCTAGGCGCTGCTCAAAGGGGCAAAAACCCGGTTGCCAT
GGCGAGGACCAAACGACAAGATGGGAAAAAAACAATAGTCTATTGTTAAATCGTAATACTGTATTGTGAGAT
GCTGACGCGTTTCGTTTTTCGTGTCAGCGTTCTTTATATTGTTTCGTGTTCTGCTGCAAAACGTATATAAAC
GCACTGCTATTTTGCCTTCTTTTGCCTTCTTCCTTGCCTTTCTCTCATCTCATATCCAAGTTGAAATAAATA
TGACAAGCAAGAATAGCAGCAGCAATAAATCAAATACTCCCACACAAATGTTAATTAACTCTAAAGGTGAAG
AATTATTCACTGGTGTTGTCCCAATTTTGGTTGAATTAGATGGTGATGTTAATGGTCACAAATTTTCTGTCT
CCGGTGAAGGTGAAGGTGATGCTACTTACGGTAAATTGACCTTAAAATTTATTTGTACTACTGGTAAATTGC
CAGTTCCATGGCCAACCTTAGTCACTACTTTCGGTTATGGTGTTCAATGTTTTGCGAGATACCCAGATCATA
TGAAACAACATGACTTTTTCAAGTCTGCCATGCCAGAAGGTTATGTTCAAGAAAGAACTATTTTTTTCAAAG
ATGACGGTAACTACAAGACCAGAGCTGAAGTCAAGTTTGAAGGTGATACCTTAGTTAATAGAATCGAATTAA
AAGGTATTGATTTTAAAGAAGATGGTAACATTTTAGGTCACAAATTGGAATACAACTATAACTCTCACAATG
TTTACATCATGGCTGACAAACAAAAGAATGGTATCAAAGTTAACTTCAAAATTAGACACAACATTGAAGATG
GTTCTGTTCAATTAGCTGACCATTATCAACAAAATACTCCAATTGGTGATGGTCCAGTCTTGTTACCAGACA
ACCATTACTTATCCACTCAATCTGCCTTATCCAAAGATCCAAACGAAAAGAGAGACCACATGGTCTTGTTAG
AATTTGTTACTGCTGCTGGTATTACCCATGGTATGGATGAATTGTACAAATAACTGCAGGGCGCGCCACTTC
TAAATAAGCGAATTTCTTATGATTTATGATTTTTATTATTAAATAAGTTATAAAAAAAATAAGTGTATACAA
ATTTTAAAGTGACTCTTAGGTTTTAAAACGAAAATTCTTATTCTTGAGTAACTCTTTCCTGTAGGTCAGGTT
GCTTTCTCAGGTATAGTATGAGGTCGCTCTTATTGACCACACCTCTACCGGCAGATCCGCTAGGGATAACAG
GGTAATATAGATCTGCCCGCCGGGAAGGCGAACCCGATCGGATGCATCCTCTCTGCTGCCATGATGCTGAAG
TTGTCGTTGAACATGGTTGCTGCCGGCGAGGCGGTCGAGCAGGCAGTGCAGGAGGTGTTGGACTCGGGAGTC
AGAACGGGCGACCTGCTCGGCTCGAGCTCCACTTCGGAGGTTGGCGACGCCATTGCGCTTGCAGTTAAGGAA
GCCTTGCGCAGGCAATCCGCAGCTGGTCTGAGCTAGCCTCGAGGACCCTTCTCTTTAGACTATTCTACTCTT
ATGCACGTAAAAAATTCTAGGAAATATGTATTAACTAGGAGTAAAATAACCGGCTAGTGGCATTCATATAGC
CGTCTGTTTACATCTACATCACACATTTCGAGTGTATATCTCGCAACGTTGGCGTTAAATAGGCAGTCAATG
GCCCGACCATTCTATGGTGTTTAGGTCGATGCCATCTTTGTACGTTTAGCTTATCGATGATAAGCTGTCAAA
CATGAGAATTCTTGAAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATG
GTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATA
CATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGT
ATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCAC
CCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGAT
CTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTT
CTGCTATGTGGCGCGGTATTATCCCGTGTTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCT
CAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTA
TGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAG
GAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAAT
GAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGCAGCAATGGCAACAACGTTGCGCAAACTATTA
ACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGA
CCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCT
CGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGT
CAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTG
TCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTG
AAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCC
GTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAA
CCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTC
AGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTA
GCACCGCCTACATACCTCGCTC

TGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGAT
AGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGA
CCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGG
ACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGT
ATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGC
GGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACA
TGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTC
GCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTC
TCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCAT
AGTTAAGCCAGTATACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACACCCGCCAACACC
CGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAG
CTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAGGCAGAGCTTTGAAGAAAAATGCGCCTT
ATTCAATCTTTGCTATAAAAAATGGCCCAAAATCTCACATTGGAAGACATTTGATGACCTCATTTCTTTCAA
TGAAGGGCCTAACGGAGTTGACTAATGTTGTGGGAAATTGGAGCGATAAGCGTGCTTCTGCCGTGGCCAGGA
CAACGTATACTCATCAGATAACAGCAATACCTGATCACTACTTCGCACTAGTTTCTCGGTACTATGCATATG
ATCCAATATCAAAGGAAATGATAGCATTGAAGGATGAGACTAATCCAATTGAGGAGTGGCAGCATATAGAAC
AGCTAAAGGGTAGTGCTGAAGGAAGCATACGATACCCCGCATGGAATGGGATAATATCACAGGAGGTACTAG
ACTACCTTTCATCCTACATAAATAGACGCATATAAGTACGCATTTAAGCATAAACACGCACTATGCCGTTCT
TCTCATGTATATATATATACAGGCAACACGCAGATATAGGTGCGACGTGAACAGTGAGCTGTATGTGCGCAG
CTCGCGTTGCATTTTCGGAAGCGCTCGTTTTCGGAAACGCTTTGAAGTTCCTATTCCGAAGTTCCTATTCTC
TAGAAAGTATAGGAACTTCAGAGCGCTTTTGAAAACCAAAAGCGCTCTGAAGACGCACTTTCAAAAAACCAA
AAACGCACCGGACTGTAACGAGCTACTAAAATATTGCGAATACCGCTTCCACAAACATTGCTCAAAAGTATC
TCTTTGCTATATATCTCTGTGCTATATCCCTATATAACCTACCCATCCACCTTTCGCTCCTTGAACTTGCAT
CTAAACTCGACCTCTACATTTTTTATGTTTATCTCTAGTATTACTCTTTAGACAAAAAAATTGTAGTAAGAA
CTATTCATAGAGTGAATCGAAAACAATACGAAAATGTAAACATTTCCTATACGTAGTATATAGAGACAAAAT
AGAAGAAACCGTTCATAATTTTCTGACCAATGAAGAATCATCAACGCTATCACTTTCTGTTCACAAAGTATG
CGCAATCCACATCGGTATAGAATATAATCGGGGATGCCTTTATCTTGAAAAAATGCACCCGCAGCTTCGCTA
GTAATCAGTAAACGCGGGAAGTGGAGTCAGGCTTTTTTTATGGAAGAGAAAATAGACACCAAAGTAGCCTTC
TTCTAACCTTAACGGACCTACAGTGCAAAAAGTTATCAAGAGACTGCATTATAGAGCGCACAAAGGAGAAAA
AAAGTAATCTAAGATGCTTTGTTAGAAAAATAGCGCTCTCGGGATGCATTTTTGTAGAACAAAAAAGAAGTA
TAGATTCTTTGTTGGTAAAATAGCGCTCTCGCGTTGCATTTCTGTTCTGTAAAAATGCAGCTCAGATTCTTT
GTTTGAAAAATTAGCGCTCTCGCGTTGCATTTTTGTTTTACAAAAATGAAGCACAGATTCTTCGTTGGTAAA
ATAGCGCTTTCGCGTTGCATTTCTGTTCTGTAAAAATGCAGCTCAGATTCTTTGTTTGAAAAATTAGCGCTC
TCGCGTTGCATTTTTGTTCTACAAAATGAAGCACAGATGCTTCGTTCTGCGGTAAAGCTCATCAGCGTGGTC
GTGAAGCGATTCACAGATGTCTGCCTGTTCATCCGCGTCCAGCTCGTTGAGTTTCTCCAGAAGCGTTAATGT
CTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTTTCCTGTTTGGTCACTGATGCCTCCGTGTAAGG
GGGATTTCTGTTCATGGGGGTAATGATACCGATGAAACGAGAGAGGATGCTCACGATACGGGTTACTGATGA
TGAACATGCCCGGTTACTGGAACGTTGTGAGGGTAAACAACTGGCGGTATGGATGCGGCGGGACCAGAGAAA
AATCACTCAGGGTCAATGCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCCAGCAGCATCC
TGCGATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTTCCAGACTTTACGAAACACGGAA
ACCGAAGACCATTCATGTTGTTGCTCAGGTCGCAGACGTTTTGCAGCAGCAGTCGCTTCACGTTCGCTCGCG
TATCGGTGATTCATTCTGCTAACCAGTAAGGCAACCCCGCCAGCCTAGCCGGGTCCTCAACGACAGGAGCAC
GATCATGCGCACCCGTGGCCAGGACCCAACGCTGCGGGGGGGGGGGGGGTTTTCTTTCCAATTTTTTTTTTT
TCGTCATTATAGAAATCATTACGACCGAGATTCCCGGGTAATAACTGATATAATTAAATTGAAGCTCTAATT
TGTGAGTTTAGTATACATGCATTTACTTATAATACAGTTTTTTAGTTTTGCTGGCCGCATCTTCTCAAATAT
GCTTCCCAGCCTGCTTTTCTGTAACGTTCACCCTCTACCTTAGCATCCCTTCCCTTTGCAAATAGTCCTCTT
CCAACAATAATAATGTCAGATCCTGTAGAGACCACATCATCCACGGTTCTATACTGTTGACCCAATGCGTCT
CCCTTGTCATCTAAACCCACACCGGGTGTCATAATCAACCAATCGTAACCTTCATCTCTTCCACCCATGTCT
CTTTGAGCAATAAAGCCGATAACAAAATCTTTGTCGCTCTTCGCAATGTCAACAGTACCCTTAGTATATTCT
CCAGTAGCTAGGGAGCCCTTGCATGACAATTCTGCTAACATCAAAAGGCCTCTAGGTTCCTTTGTTACTTCT
TCCGCCGCCTGCTTCAAACCGCTAACAATACCTGGGCCCACCACACCGTGTGCATTCGTAATGTCTGCCCAT
TCTGCTATTCTGTATACACCCGCAGAGTACTGCAATTTGACTGTATTACCAATGTCAGCAAATTTTCTGTCT
TCGAAGAGTAAAAAATTGTACTTGGCGGATAATGCCTTTAGCGGCTTAACTGTGCCCTCCATGGAAAAATCA
GTCAAGATATCCACATGT

GTTTTTAGTAAACAAATTTTGGGACCTAATGCTTCAACTAACTCCAGTAATTCCTTGGTGGTACGAACATCC
AATGAAGCACACAAGTTTGTTTGCTTTTCGTGCATGATATTAAATAGCTTGGCAGCAACAGGACTAGGATGA
GTAGCAGCACGTTCCTTATATGTAGCTTTCGACATGATTTATCTTCGTTTCCTGCAGGTTTTTGTTCTGTGC
AGTTGGGTTAAGAATACTGGGCAATTTCATGTTTCTTCAACACCACATATGCGTATATATACCAATCTAAGT
CTGTGCTCCTTCCTTCGTTCTTCCTTCTGCTCGGAGATTACCGAATCAAAAAAATTTCAAAGAAACCGGAAT
CAAAAAAAAGAACAAAAAAAAAAAAGATGAATTGAAACCCCCCCCCCCCCCGATGCGCCGCGTGCGGCTGCT
GGAGATGGCGGACGCGATGGATATGTTCTGCCAAGGGTTGGTTTGCGCATTCACAGTTCTCCGCAAGAATTG
ATTGGCTCCAATTCTTGGAGTGGTGAATCCGTTAGCGAGGTGCCGCCGGCTTCCATTCAGGTCGAGGTGGCC
CGGCTCCATGCACCGCGACGCAACGCGGGGAGGCAGACAAGGTATAGGGCGGCGCCTACAATCCATGCCAAC
CCGTTCCATGTGCTCGCCGAGGCGGCATAAATCGCCGTGACGATCAGCGGTCCAGTGATCGAAGTTAGGCTG
GTAAGAGCCGCGAGCGATCCTTGAAGCTGTCCCTGATGGTCGTCATCTACCTGCCTGGACAGCATGGCCTGC
AACGCGGGCATCCCGATGCCGCCGGAAGCGAGAAGAATCATAATGGGGAAGGCCATCCAGCCTCGCGTCGCG
AACGCCAGCAAGACGTAGCCCAGCGCGTCGGCCGCCATGCCGGCGATAATGGCCTGCTTCTCGCCGAAACGT
TTGGTGGCGGGACCAGTGACGAAGGCTTGAGCGAGGGCGTGCAAGATTCCGAATACCGCAAGCGACAGGCCG
ATCATCGTCGCGCTCCAGCGAAAGCGGTCCTCGCCGAAAATGACCCAGAGCGCTGCCGGCACCTGTCCTACG
AGTTGCATGATAAAGAAGACAGTCATAAGTGCGGCGACGATAGTCATGCCCCGCGCCCACCGGAAGGAGCTG
ACTGGGTTGAAGGCTCTCAAGGGCATCGGTCGACGCTCTCCCTTATGCGACTCCTGCATTAGGAAGCAGCCC
AGTAGTAGGTTGAGGCCGTTGAGCACCGCCGCCGCAAGGAATGGTGCATGCAAGGAGATGGCGCCCAACAGT
CCCCCGGCCACGGGGCCTGCCACCATACCCACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCCCGA
TCTTCCCCATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACG
ATGCGTCCGGCGTAGA

FIG. 28

A


RNR2-GFP +/- MMS (n=5)
Brightness
100
90
80
70
60
50
40
30
20
10
0
RNR2 - MMS
RNR2 +0.005%
RNR2 +0.01%

B

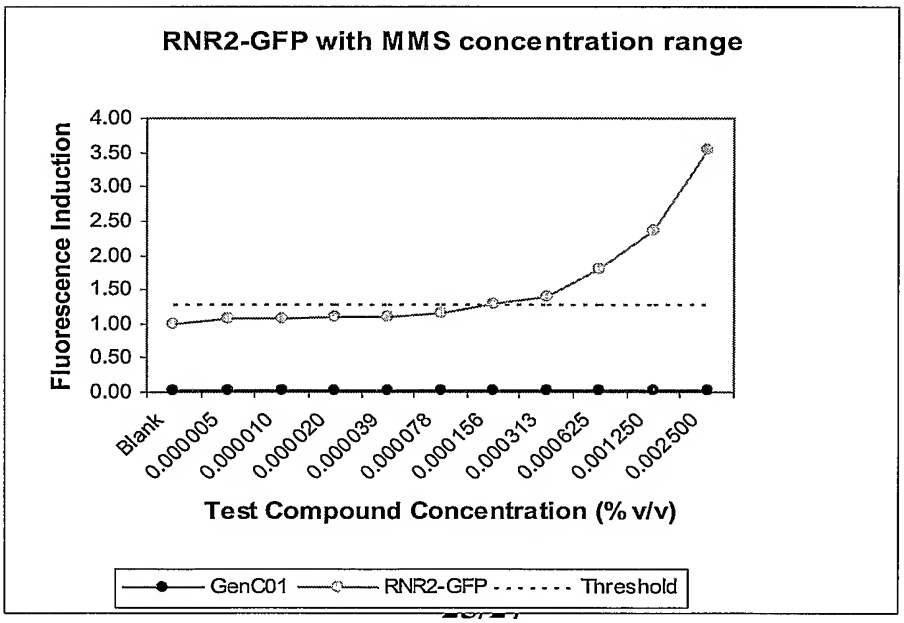
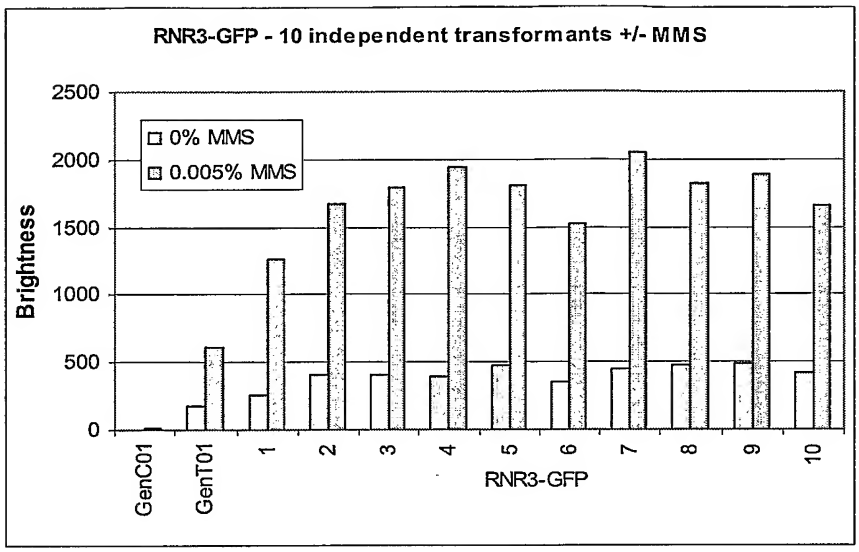

RNR2-GFP with MMS concentration range
Fluorescence Induction
4.00
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
Blank
0.000005
0.000010
0.000020
0.000039
0.000078
0.000156
0.000313
0.000625
0.001250
0.002500
Test Compound Concentration (% v/v)
GenC01
RNR2-GFP
Threshold

FIG. 29

A


RNR3-GFP - 10 independent transformants +/- MMS
Brightness
2500
2000
1500
1000
500
0
0% MMS
0.005% MMS
GenC01
GenT01
1
2
3
4
5
6
7
8
9
10
RNR3-GFP

B

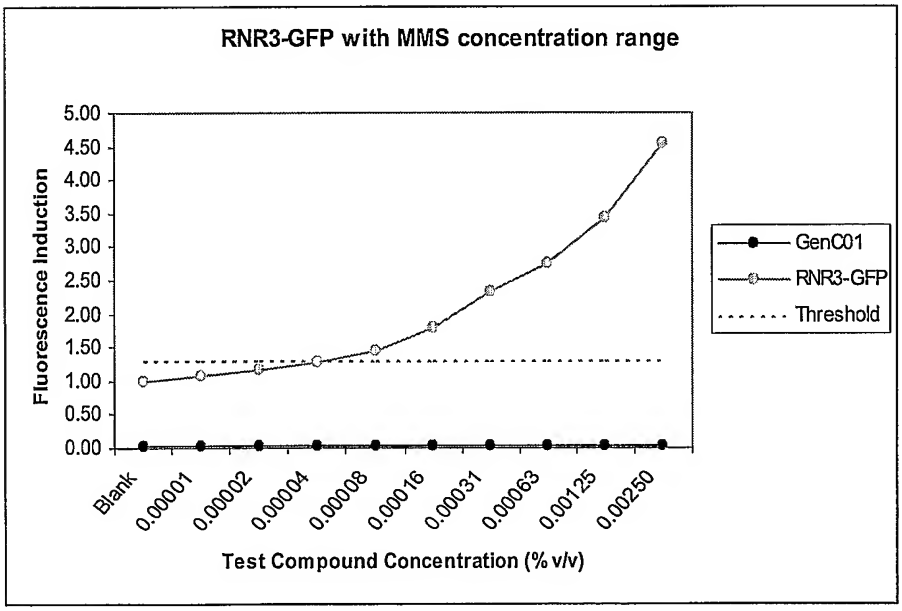
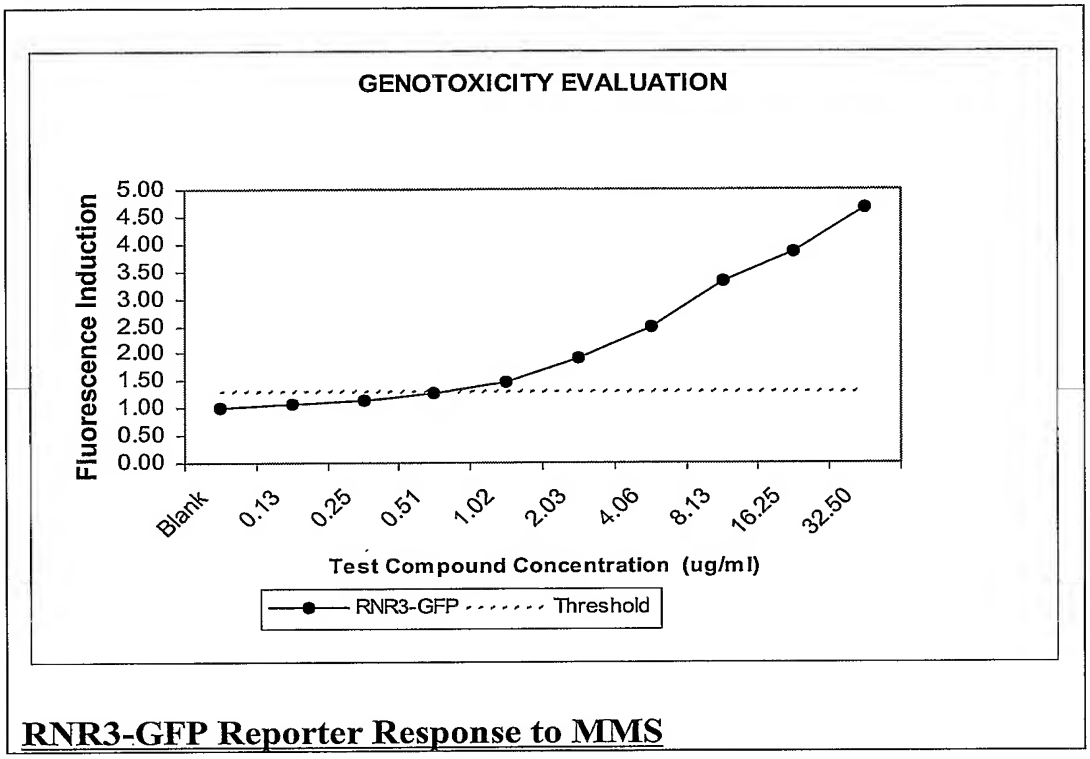

RNR3-GFP with MMS concentration range
Fluorescence Induction
5.00
4.50
4.00
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
GenC01
RNR3-GFP
Threshold
Blank
0.00001
0.00002
0.00004
0.00008
0.00016
0.00031
0.00063
0.00125
0.00250
Test Compound Concentration (% v/v)

Fig:30

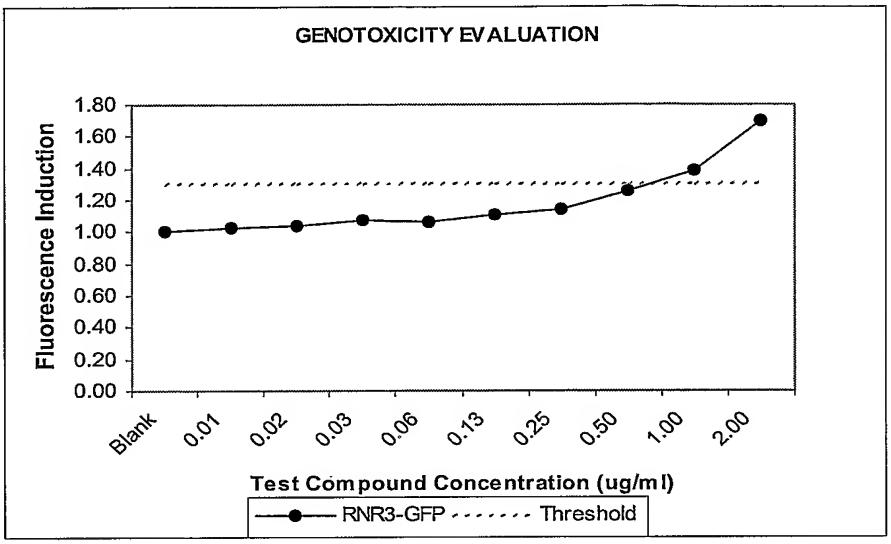
Methyl methanesulfonate



RNR3-GFP Reporter Response to MMS

Fig:31

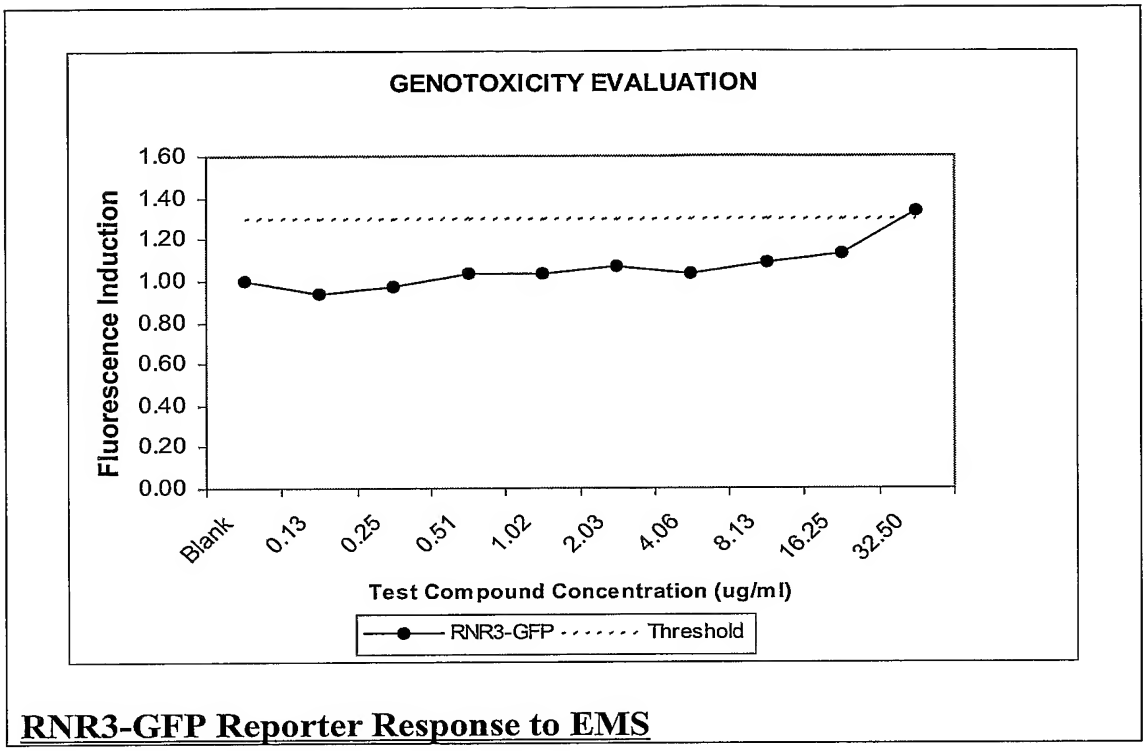
1,2-Dimethylhydrazine (dihydrochloride)



RNR3-GFP Reporter Response to 1,2-Dimethylhydrazine (2HCl)

Fig:32

Ethyl methanesulfonate



RNR3-GFP Reporter Response to EMS

Fig:33

RNR3 sequence data downloaded from SGD, Chromosome IX:

```
   1  AGATTCTGCG CCAGCAAGTC GTCTCCGAGG GGGGGCCCAC CGCTACCATA
  51  CAGTCAAGGT TTAACTACGC ATGGGGGCTG ATCAAATCCA CTGACGTGAA
 101  TGACGAAAGG CTTGGTGTGA AAATCCTCAC AGACATTTAC AAAGAGGCCG
 151  AGTCCCGTAG ACGAGAATGC CTATATTATC TGACCATAGG TTGCTACAAA
 201  CTCGGTGAAT ACTCTATGGC GAAGAGATAT GTAGACACTT TATTTGAGCA
 251  TGAGCGTAAT AACAAGCAGG TGGGCGCTTT GAAGAGTATG GTAGAGGATA
 301  AGATCCAGAA GGAAACACTC AAGGGTGTTG TCGTCGCTGG AGGCGTACTA
 351  GCCGGCGCTG TGGCCGTGGC TAGTTTCTTC TTAAGAAACA AGAGAAGGTA
 401  ACAAGCACAT AAAAAATCAG CACATACGTA CATACATAAG AATGAATCGC
 451  ACGCACGCGT AAACATTTAT CATTTAATCT TCAGTTGTTA GATAAAAAAA
 501  AAAAGAAAAG AAAAGAAAGT GAAGGCTTGT TTCAGTTTGA ACTAGGTAGC
 551  AGAGCAAGCC CTCGTTCTTG GCTGCTAATT TTCCTAAAGT AGTAAAAAAA
 601  GCCAAGTTAT CTGCCTACGG TTGTCACAGC AACATTGCGT GCCGTTGTTC
 651  TTTTGTTTTT TTTTTTTTTT TTTTTTCGTG GTTGTCGCAG CAACGACACC
 701  TAGGCGCTGC TCAAAGGGGC AAAAACCCGG TTGCCATGGC GAGGACCAAA
 751  CGACAAGATG GGAAAAAAAC AATAGTCTAT TGTTAAATCG TAATACTGTA
 801  TTGTGAGATG CTGACGCGTT TCGTTTTTCG TGTCAGCGTT CTTTATATTG
 851  TTTCGTGTTC TGCTGCAAAA CGTATATAAA CGCACTGCTA TTTTGCCTTC
 901  TTTTGCCTTC TTCCTTGCCT TTCTCTCATC TCATATCCAA GTTGAAATAA
 951  ATATGACAAG CAAGAATAGC AGCAGCAATA AATCAAATAC TCCCACACAA
1001  ATGTACGTTA TTAAAAGAGA CGGCCGCAAA GAGCCCGTTC AATTCGATAA
1051  AATTACCTCC CGTATCACCC GTTTGTCATA CGGTTTAGAC CCAAACCGTA
1101  TTGATGCTGT TAAGGTAACC CAACGTATTA TTTCTGGTGT GTACTCCGGT
1151  GTTACTACCG TTGAGCTGGA CAATCTTGCA GCTGAAACAT GTGCATACAT
1201  GACCACTGTG CACCCTGATT ATGCCACTCT AGCCGCTAGA ATCGCCATCT
1251  CTAACTTACA TAAGCAAACC ACAAAGCAAT TCTCCAAAGT TATTGAGGAT
1301  TTACACGACT GGATTAACCC AGCTACTGGA AAGCATGCTC CTATGATTTC
1351  GGACGAAATT TACAACATTG TCATGGAAAA CAAAGATACT TTGAACTCGG
1401  CCATCGTGTA CGATAGGGAT TTCCAGTATA CGTATTTCGG ATTCAAGACA
1451  CTGGAGCGTT CGTACTTGCT AAGACTGAAC GGTGAAGTGG CAGAACGTCC
1501  TCAGCATTTG GTAATGCGTG TGGCGCTAGG TATCCATGGT AGCGATATCG
1551  AATCTGTGCT GAAGACTTAT AATTTGATGT CGTTAAGATA CTTCACTCAC
1601  GCTTCCCCAA CTTTATTCAA CGCTGGTACG CCACATCCTC AAATGTCTTC
1651  ATGTTTCTTA ATTGCCATGA AGGATGACTC TATCGAAGGT ATTTATGATA
1701  CTTTGAAAGA ATGTGCTATG ATTTCCAAAA CTGCAGGTGG TGTTGGTCTT
1751  CATATCAACA ACATCCGTTC CACAGGTTCT TATATCGCTG GTACCAACGG
1801  TACTTCAAAC GGGTTGATTC CTATGATTCG TGTTTTCAAT AATACTGCCC
1851  GTTATGTGGA CCAGGGTGGT AACAAGAGAC CTGGTGCTTT CGCCCTTTTC
1901  TTGGAGCCAT GGCATGCAGA TATCTTCGAC TTTGTCGATA TCAGAAAAAC
1951  ACATGGTAAG GAAGAAATTC GTGCAAGAGA TTTGTTCCCT GCTCTATGGA
2001  TCCCTGATCT TTTCATGAAA CGTGTTCAAG AGGATGGGCC TTGGACTTTG
2051  TTTTCGCCCA GTGCTGCCCC AGGTTTAGAT GATGTGTGGG GTGATGAATT
```

2101 TGAAGAACTA TATACTCGTT ACGAAAGAGA AGGTCGTGGT AAAACAATTA
2151 AAGCCCAAAA GTTGTGGTAT GCCATTTTGC AAGCACAGAC AGAAACAGGT
2201 ACACCTTTCA TGGTTTATAA GGACGCATGT AACAGGAAGA CAAACCAACA
2251 GAACTTAGGT ACTATCAAAT CATCTAATTT ATGTTGTGAA ATCGTCGAAT
2301 ATTCCTCCCC GGATGAAACT GCAGTTTGTA ATTTAGCTTC TATTGCCCTA
2351 CCAGCATTCG TTGAGGTTTC AGAAGATGGT AAAACTGCAA GCTATAATTT
2401 CGAGAGATTA CACGAGATTG CTAAAGTCAT TACTCACAAC TTGAACAGAG
2451 TTATCGACCG TAATTACTAT CCAGTTCCCG AGGCTAGAAA TTCAAATATG
2501 AAGCATAGAC CTATTGCTCT TGGTGTCCAG GGTTTGGCCG ATACTTATAT
2551 GATGTTGCGT CTACCCTTTG AATCGGAAGA AGCTCAAACT CTAAACAAAC
2601 AAATCTTCGA AACTATTTAC CATGCTACTC TTGAAGCCTC CTGTGAATTG
2651 GCCCAAAAAG AAGGTAAATA TTCTACTTTT GAAGGTTCTC CAGCTTCTAA
2701 GGGTATTTTA CAATTCGATA TGTGGAACGC TAAACCATTT GGCATGTGGG
2751 ATTGGGAAAC CTTAAGAAAG GACATTGTTA AACATGGGTT AAGAAACTCT
2801 TTGACTATGG CACCAATGCC AACCGCCTCA ACTTCCCAAA TTCTTGGTTA
2851 TAATGAATGC TTCGAACCAG TGACCTCAAA CATGTACTCT CGTCGTGTCC
2901 TGTCTGGTGA ATTCCAAGTT GTTAATCCAT ATTTACTACG TGATTTAGTC
2951 GACCTGGGTA TTTGGGATGA TAGTATGAAA CAATATCTAA TTACACAAAA
3001 TGGTTCTATT CAAGGCTTAC CAAATGTGCC ACAAGAATTG AAGGAATTAT
3051 ACAAAACCGT CTGGGAAATC TCTCAAAAGA CCATTATCAA TATGGCTGCT
3101 GATCGTGCCA TCTACATCGA TCAGTCTCAT TCCTTGAATC TTTTCTTGCA
3151 AGCACCATCA ATGGGTAAGA TTACTAGTAT GCATTTCTAC GGTTGGAAGA
3201 AGGGTTTAAA AACTGGTATG TACTACTTAA GAACGCAAGC CGCCTCCGCT
3251 GCTATTCAAT TTACCATTGA TCAAGAGGTT GCCGATCAAG CCGCTACACA
3301 TATTGCTTCC GTCTCAGAAT TGGATCGTCC AGTTTATGTT CCAAAGGGTA
3351 CAAAATTCTC TGAACAAAAG GCGGCATCTG CGCTTACCGA AAGCTCAGAT
3401 AATGAGAAGG ATGCATCTCC AGTTCCATCC GAACAATCAT CGGTGTCGAG
3451 TGCCATGTCA AATGTGAAAT TGGAAGATAG TGTTGCCCCA GCAGTTCCAA
3501 CGGAAACAAT AAAAGAAGAT TCCGACGAGA AGAAATGTGA CATTTACAAT
3551 GAAAAGGTGA TTGCTTGTAC TGCTCCTACT CCAGAAGCTT GTGAGTCATG
3601 TTCCGGTTGA

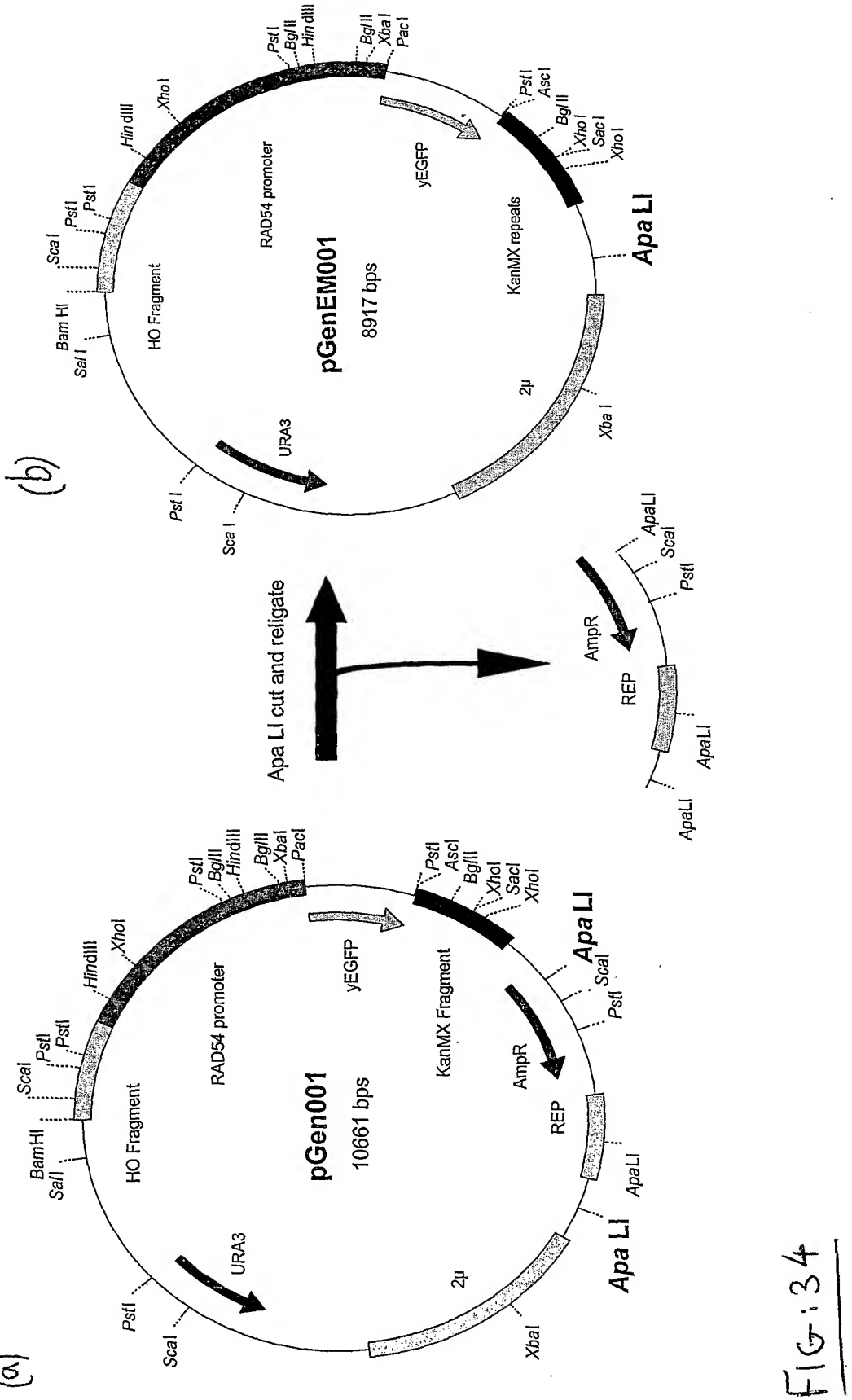

Removal of Bacterial Origin of replication and Amp Resistance
(a)
pGen001
10661 bps
HO Fragment
RAD54 promoter
yEGFP
KanMX Fragment
AmpR
REP
2μ
URA3
Apa LI cut and religate
(b)
pGenEM001
8917 bps
HO Fragment
RAD54 promoter
yEGFP
KanMX repeats
2μ
URA3
Apa LI
AmpR
REP

Fig: 34

Fig:35

Fragment of HO sequence used in the integrating vector (pWDH443)

```
   1  AAATTGTGAC AGCTTTCCAG AATGGATTAT TTTTCCTCAA ATTCCTTGTC
  51  TTCCTGTTTT CATCTGGACC ATCTCCATAA TGAAGCCTTA CATGTTTGGC
 101  ACGTAGCGGA ACGTGATCGT CACAAACCGT AAGGTAGAGA CCCCAGATTT
 151  TCGCATTTTC TCTTAAACTC TCCATTAGCT TAGGATCCAA GCTATCTACT
 201  GAGATTTCTG GCTCTTTTGT TGTACTGTCA CCTAACCACA GACCAAGCAT
 251  CCAAGCCATA CTTTTTACAG CAGGAGTTAC AAGGTCACTA CGTCCAGTGA
 301  GAAATTTAGA TAAAACACCA TTTCCTGCGA GTACTGGACC AAATCTTATG
 351  CAGCTAGAAA TTCTCAATTG AGCATCAAGA TAATCCAAAT CTCTAACTTC
 401  AATGTCAAAG TTGAAATATT CTCCTTTAGA GCGCTCCATT TCTTCTATGA
 451  AGCGTTTTGC GGCAAACTCA CCTTCAACTG TCATTGGGAA TGTCTTATGA
 501  TGGTTTTTTG GAATTATTAT TATCCTACCA TCAAGCGTCT GACATTGCTG
 551  CAGATTTCTC CATCTCACTT TATATTTGGT GGCATTTCTA CCACTTTTTT
 601  CCAACAGTGG TTTGGTAGGG ACCCTGACTG ACAATTTATG ACCTGCAGTA
 651  CATTGTAATG CAAGACGCTG ATAAACTGTT CTACGCCTGG GATCTAACCT
 701  ACCAGGTTCA CCTTCAAAAG CTCTGTGTTT GGTTTTTTGC TGTATATTAT
 751  AGATTTTCTG ATAGCCCTGT GTGACATTTA TGACGCGGGC AGCGGAGCCA
 801  TCTGCGCACA TAACGTAAGA GTTAGCCGTG ACGTTTGCGA TGTCTTTAAT
 851  TTCACCGTTA GCCATCAGAA TAGTCGTGTT TTCAGAAAGC AT
```

Fig:36

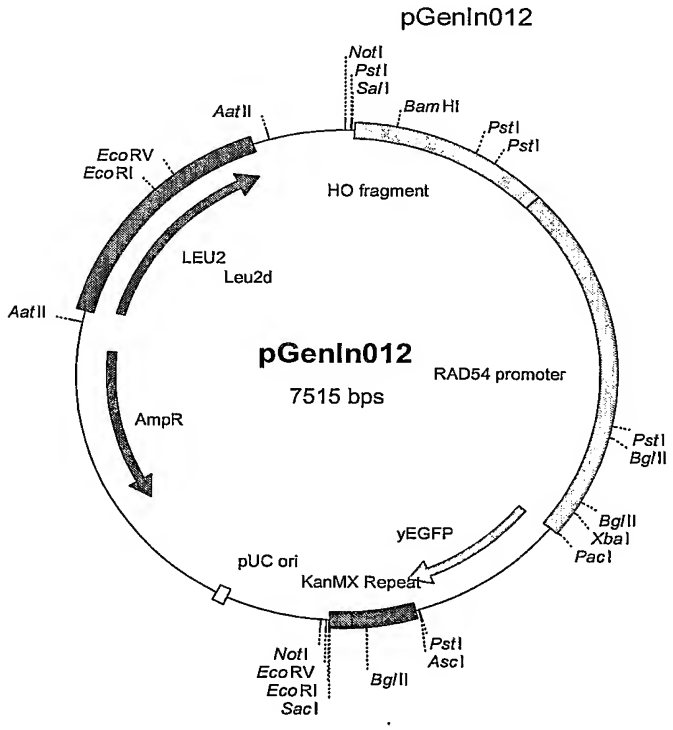


Fig:37

rDNA sequence used in multiple copy rDNA integrating plasmids

```
   1  GATCTGACGA TCACCTAGCG ACTCTCTCCA CCGTTTGACG AGGCCATTTA
  51  CAAAAACATA ACGAACGACA AGCCTACTCG AATTCGTTTC CAAACTCTTT
 101  TCGAACTTGT CTTCAACTGC TTTCGCATGA AGTACCTCCC AACTACTTTT
 151  CCTCACACTT GTACTCCATG ACTAAACCCC CCCTCCCATT ACAAACTAAA
 201  ATCTTACTTT TATTTTCTTT TGCCCTCTCT GTCGCTCTGC CTTAACTACG
 251  TATTTCTCGC CGAGAAAAAC TTCAATTTAA GCTATTCTCC AAAAATCTTA
 301  GCGTATATTT TTTTTCCAAA GTGACAGGTG CCCCGGGTAA CCCAGTTCCT
 351  CACTATTTTT TACTGCGGAA GCGGAAGCGG AAAATACGGA AACGCGCGGG
 401  AACATACAAA ACATACAAAA TATACCTTTC TCACACAAGA AATATATGCT
 451  ACTTGCAAAA TATCATACCA AAAAACTTTT CACAACCGAA ACCAAAACCA
 501  ACGGATATCA TACATTACAC TACCACCATT CAAACTTTAC TACTATCCTC
 551  CCTTCAGTTT CCCTTTTTCT GCCTTTTTCG GTGACGGAAA TACGCTTCAG
 601  AGACCCTAAA GGGAAATCCA TGCCATAACA GGAAAGTAAC ATCCCAATGC
 651  GGACTATACC ACCCCACCAC ACTCCTACCA ATAACGGTAA CTATTCTATG
 701  TTTTCTTACT CCTATGTCTA TTCATCTTTC ATCTGACTAC CTAATACTAT
 751  GCAAAAATGT AAAATCATCA CACAAAACAT AAACAATCAA AATCAGCCAT
 801  TTCCGCACCT TTTCCTCTGT CCACTTTCAA CCGTCCCTCC AAATGTAAAA
 851  TGGCCTATCG GAATACATTT TCTACATCCT AACTACTATA AAACAACCTT
 901  TAGACTTACG TTTGCTACTC TCATGGTCTC AATACTGCCG CCGACATTCT
 951  GTCCCACATA CTAAATCTCT TCCCGTCATT ATCGCCCGCA TCCGGTGCCG
1001  TAAATGCAAA ACAAATACCA TCTATGTCTT CCACACCATC ATTTTACTAT
1051  GCCTGCCACC ATCCATTTGT CTTTTGCACC ATATCTTCAT AACCTGTCAC
1101  CTTGAAACTA CCTCTGCATG CCACCTACCG ACCAACTTTC ATGTTCTGTT
1151  TCGACCTACC TCTTGTAAAT GACAAATCAC CTTTTTCATC GTATGCACCT
1201  TATTCTCCAC ATCACAATGC ACTATTGCTT TTGCTTTTTC ACCTGTCATA
1251  TCCTATTGCT ATTAGATGAA ATATAATAAA AATTGTCCTC CACCCATAAC
1301  ACCTCTCACT CCCACCTACT GAACATGTCT GGACCCTGCC CTCATATCAC
1351  CTGCGTTTCC GTTAAACTAT CGGTTGCGGC CATATCTACC AGAAAGCACC
1401  GTTTCCCGTC CGATCAACTG TAGTTAAGCT GGTAAGAGCC TGACCGAGTA
1451  GTGTAGTGGG TGACCATACG CGAAACTCAG GTGCTGCAAT CTTTATTTCT
1501  TTTTTTTTTT TTTTTTTTTT TTTTTTTTCT AGTTTCTTGG CTTCCTATGC
1551  TAAATCCCAT AACTAACCTA CCATTCGATT CAGAAAAATT CGCACTATCC
1601  AGCTGCACTC TTCTTCTGAA GAGTTAAGCA CTCCATTATG CTCATTGGGT
1651  TGCTACTACT TGATATGTAC AAACAATATT CTCCTCCGAT ATTCCTACAA
1701  AAAAAAAAAA AAAAACACTC CGGTTTTGTT CTCTTCCCTC CATTTCCCTC
1751  TCTTCTACGG TTAATACTTT CCTCTTCGTC TTTTTCTACA CCCTCGTTTA
1801  GTTGCTTCTT ATTCCTTCCC GCTTTCCTGC ACTAACATTT TGCCGCATTA
1851  CACTATATGA TCGTAGTACA TCTTACAACT CCGCATACCG CGTCGCCGCG
1901  TCGCCGCGTC GCCAAAAATT TACTTCGCCA ACCATTCCAT ATCTGTTAAG
1951  TATACATGTA TATATTGCAC TGGCTATTCA TCTTGCACTT TTCCTCTTTC
2001  TTCTTCCCAG TAGCCTCATC CTTTTACGCT GCCTCTCTGG AACTTGCCAT
2051  CATCATTCCC TAGAAACTGC CATTTACTTA AAAAAAAAAA AAAAAAAAAA
2101  ATGTCCCCAC TGTTCACTGT TCACTGTTCA CTTGTCTCTT ACATCTTTCT
2151  TGGTAAAATC GTAGTTCGTA GTATTTTTTT TCATATCAAA GGCATGTCCT
2201  GTTAACTATA GGAAATGAGC TTTTCTCAAT TCTCTAAACT TATACAAGCA
2251  CCTCATGTTT GCCGCTCTGA TGGTGCGGAA AAAACTGCTC CATGAAGCAA
2301  ACTGTCCGGG CAAATCCTTT CACGCTCGGG AAGCTTTGTG AAAGCCCTTC
2351  TCTTTCAACC CATCTTTGCA ACGAAAAAAA AAAAAAAAAT AAAAAATAAA
```

2401 AAGACCAAAT AGTAAATAGT AACTTACATA CATTAGTAAA TGGTACACTC
2451 TTACACACTA TCATCCTCAT CGTATATTAT AATAGATATA TACAATACAT
2501 GTTTTTACCC GGATCATAGA ATTCTTAAGA CAAATAAAAT TTATAGAGAC
2551 TTGTTCAGTC TACTTCTCTC TAAACTAGGC CCCGGCTCCT GCCAGTACCC
2601 ACTTAGAAAG AAATAAAAAA CAAATCAGAC AACAAAGGCT TAATCTCAGC
2651 AGATCGTAAC AACAAGGCTA CTCTACTGCT TACAATACCC CGTTGTACAT
2701 CTAAGTCGTA TACAAATGAT TTATCCCCAC GCAAAATGAC ATTGCAATTC
2751 GCCAGCAAGC ACCCAAGGCC TTTCCGCCAA GTGCACCGTT GCTAGCCTGC
2801 TATGGTTCAG CGACGCCACA AGGACGCCTT ATTCGTATCC ATCTATATTG
2851 TGTGGAGCAA AGAAATCACC GCGTTCTAGC ATGGATTCTG ACTTAGAGGC
2901 GTTCAGCCAT AATCCAGCGG ATGGTAGCTT CGCGGCAATG CCTGATCAGA
2951 CAGCCGCAAA AACCAATTAT CCGAATGAAC TGTTCCTCTC GTACTAAGTT
3001 CAATTACTAT TGCGGTAACA TTCATCAGTA GGGTAAAACT AACCTGTCTC
3051 ACGACGGTCT AAACCCAGCT CACGTTCCCT ATTAGTGGGT GAACAATCCA
3101 ACGCTTACCG AATTCTGCTT CGGTATGATA GGAAGAGCCG ACATCGAAGA
3151 ATCAAAAAGC AATGTCGCTA TGAACGCTTG ACTGCCACAA GCCAGTTATC
3201 CCTGTGGTAA CTTTTCTGGC ACCTCTAGCC TCAAATTCCG AGGGACTAAA
3251 GGATCGATAG GCCACACTTT CATGGTTTGT ATTCACACTG AAAATCAAAA
3301 TCAAGGGGGC TTTTACCCTT TTGTTCTACT GGAGATTTCT GTTCTCCATG
3351 AGCCCCCCTT AGGACATCTG CGTTATCGTT TAACAGATGT GCCGCCCCAG
3401 CCAAACTCCC CACCTGACAA TGTCTTCAAC CCGGATCAGC CCCGAATGGG
3451 ACCTTGAATG CTAGAACGTG GAAAATGAAT TCCAGCTCCG CTTCATTGAA
3501 TAAGTAAAGA AACTATAAAG GTAGTGGTAT TTCACTGGCG CCGAAGCTCC
3551 CACTTATTCT ACACCCTCTA TGTCTCTTCA CAATGTCAAA CTAGAGTCAA
3601 GCTCAACAGG GTCTTCTTTC CCCGCTGATT CTGCCAAGCC CGTTCCCTTG
3651 GCTGTGGTTT CGCTAGATAG TAGATAGGGA CAGTGGGAAT CTCGTTAATC
3701 CATTCATGCG CGTCACTAAT TAGATGACGA GGCATTTGGC TACCTTAAGA
3751 GAGTCATAGT TACTCCCGCC GTTTACCCGC GCTTGGTTGA ATTTCTTCAC
3801 TTTGACATTC AGAGCACTGG GCAGAAATCA CATTGCGTCA ACATCACTTT
3851 CTGACCATCG CAATGCTATG TTTTAATTAG ACAGTCAGAT TCCCCTTGTC
3901 CGTACCAGTT CTAAGTTGAT CGTTAATTGT AGCAAGCGAC GGTCTACAAG
3951 AGACCTACCA AGGCCGTCTA CAACAAGGCA CGCAAGTAGT CCGCCTAGCA
4001 GAGCAAGCCC CACCAAGCAG TCCACAAGCA CGCCCGCTGC GTCTGACCAA
4051 GGCCCTCACT ACCCGACCCT TAGAGCCAAT CCTTATCCCG AAGTTACGGA
4101 TCTATTTTGC CGACTTCCCT TATCTACATT ATTCTATCAA CTAGAGGCTG
4151 TTCACCTTGG AGACCTGCTG CGGTTATCAG TACGACCTGG CATGAAAACT
4201 ATTCCTTCCT GTGGATTTTC ACGGGCCGTC ACAAGCGCAC CGGAGCCAGC
4251 AAAGGTGCTG GCCTCTTCCA GCCATAAGAC CCCATCTCCG GATAAACCAA
4301 TTCCGGGGTG ATAAGCTGTT AAGAAGAAAA GATAACTCCT CCCAGGGCTC
4351 GCGCCGACGT CTCCACATTC AGTTACGTTA CCGTGAAGAA TCCATATCCA
4401 GGTTCCGGAA TCTTAACCGG ATTCCCTTTC GATGGTGGCC TGCATAAAAT
4451 CAGGCCTTTG AAACGGAGCT TCCCCATCTC TTAGGATCGA CTAACCCACG
4501 TCCAACTGCT GTTGACGTGG AACCTTTCCC CACTTCAGTC TTCAAAGTTC
4551 TCATTTGAAT ATTTGCTACT ACCACCAAGA TC

Fig:38

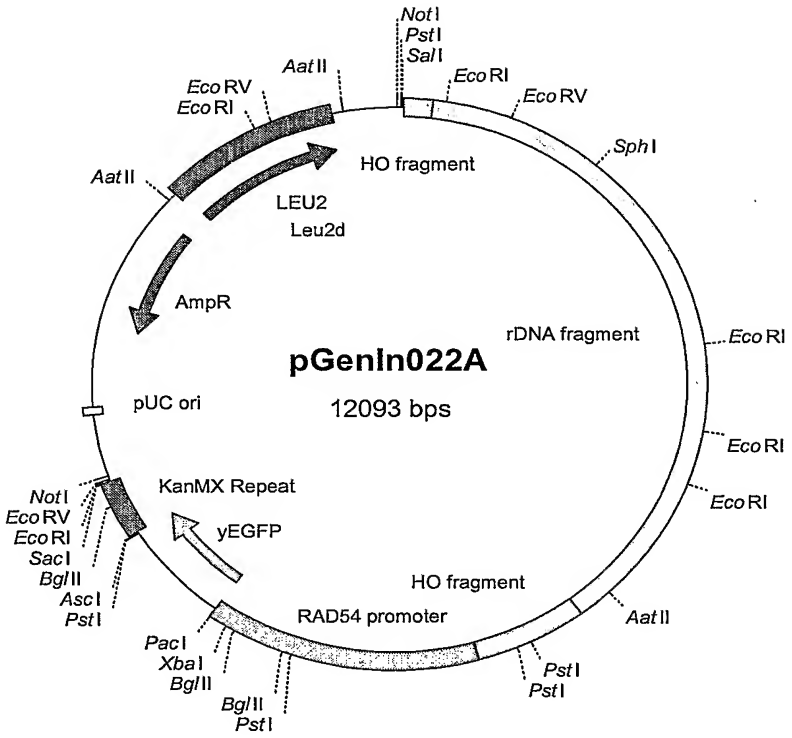

Not I
Pst I
Sal I
Aat II
Eco RV
Eco RI
Eco RI
Eco RV
Sph I
Aat II
HO fragment
LEU2
Leu2d
AmpR
rDNA fragment
Eco RI
pGenIn022A
12093 bps
pUC ori
Eco RI
Eco RI
KanMX Repeat
Not I
Eco RV
Eco RI
Sac I
Bgl II
Asc I
Pst I
yEGFP
HO fragment
RAD54 promoter
Aat II
Pac I
Xba I
Bgl II
Bgl II
Pst I
Pst I
Pst I

Fig:39

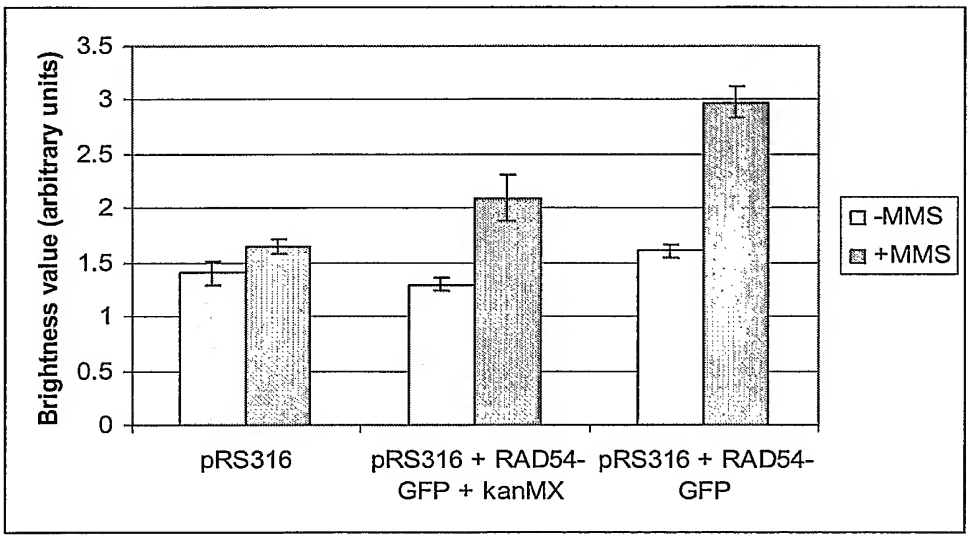

Brightness value (arbitrary units)
3.5
3
2.5
2
1.5
1
0.5
0
pRS316
pRS316 + RAD54-GFP + kanMX
pRS316 + RAD54-GFP
-MMS
+MMS

Fig:40

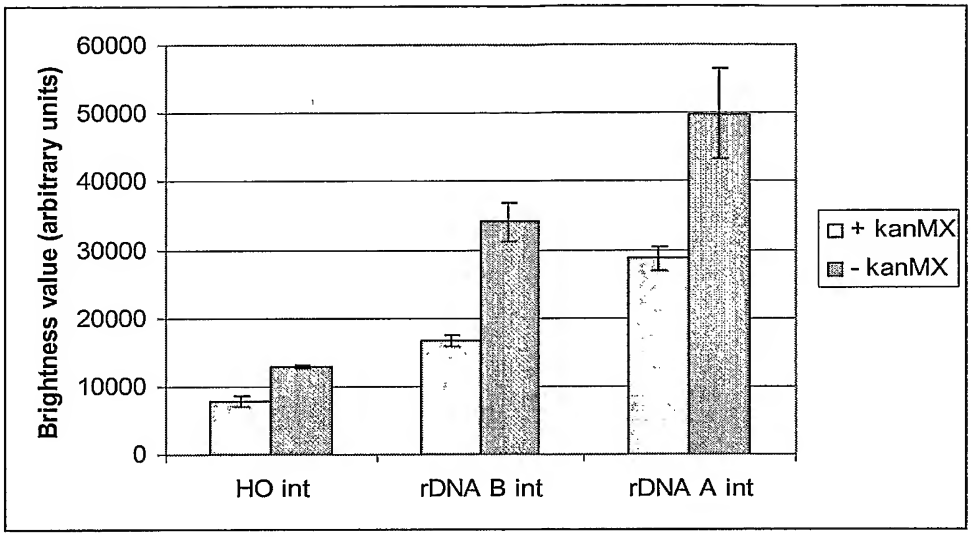

60000
50000
40000
30000
20000
10000
0
Brightness value (arbitrary units)
HO int
rDNA B int
rDNA A int
+ kanMX
- kanMX

Fig:41 pGenIn012 - 7515 bp

```
   1  GAACGCGGCC GCCAGCTGAA GCTTCGTACG CTGCAGGTCG ACGGATCAAA
  51  ATTGTGACAG CTTTCCAGAA TGGATTATTT TTCCTCAAAT TCCTTGTCTT
 101  CCTGTTTTCA TCTGGACCAT CTCCATAATG AAGCCTTACA TGTTTGGCAC
 151  GTAGCGGAAC GTGATCGTCA CAAACCGTAA GGTAGAGACC CCAGATTTTC
 201  GCATTTTCTC TTAAACTCTC CATTAGCTTA GGATCCAAGC TATCTACTGA
 251  GATTTCTGGC TCTTTTGTTG TACTGTCACC TAACCACAGA CCAAGCATCC
 301  AAGCCATACT TTTTACAGCA GGAGTTACAA GGTCACTACG TCCAGTGAGA
 351  AATTTAGATA AAACACCATT TCCTGCGAGT ACTGGACCAA ATCTTATGCA
 401  GCTAGAAATT CTCAATTGAG CATCAAGATA ATCCAAATCT CTAACTTCAA
 451  TGTCAAAGTT GAAATATTCT CCTTTAGAGC GCTCCATTTC TTCTATGAAG
 501  CGTTTTGCGG CAAACTCACC TTCAACTGTC ATTGGGAATG TCTTATGATG
 551  GTTTTTTGGA ATTATTATTA TCCTACCATC AAGCGTCTGA CATTGCTGCA
 601  GATTTCTCCA TCTCACTTTA TATTTGGTGG CATTTCTACC ACTTTTTTCC
 651  AACAGTGGTT TGGTAGGGAC CCTGACTGAC AATTTATGAC CTGCAGTACA
 701  TTGTAATGCA AGACGCTGAT AAACTGTTCT ACGCCTGGGA TCTAACCTAC
 751  CAGGTTCACC TTCAAAAGCT CTGTGTTTGG TTTTTTGCTG TATATTATAG
 801  ATTTTCTGAT AGCCCTGTGT GACATTTATG ACGCGGGCAG CGGAGCCATC
 851  TGCGCACATA ACGTAAGAGT TAGCCGTGAC GTTTGCGATG TCTTTAATTT
 901  CACCGTTAGC CATCAGAATA GTCGTGTTTT CAGAAAGCAT TTTGATCCGA
 951  CATACGATGA CCTCAATGAT TTAGATTATG TGTTGCACTT TTATAGACCT
1001  ACCAAAAATC CAGTGCGTAC ACTAATACTT TCATAAAGAT ACCTGAAACA
1051  ATAACCAGAA AGATCGGCAA AAAAATTTTT TTTCTTTGCC GAGATCACAA
1101  ACCTACTATG ACGAAAAAGC TTGAAGTTTA GATGAGTAAG GAAAATACAA
1151  GTGACGCTTT TATATGGTGC AAGGAACAAA AACTAAAAAC AACAAGGCAA
1201  ATGTGGATCT GTCATGTATG GCAACGACAG CAGGATGGCT CACAAAAAAA
1251  GACAAAAAAA ACTAAGGCAA AAGAACAAAG CTCCTCTCCT GCTCAAGAAA
1301  CGTATTGTTG AAAAACCACC GTCGTAAGAA AGTTTTTCTG TGACCTATAA
1351  TGGTTTAAAA TCGGCCCATT TTTTTTCCCT CTTTTGTGGT CCAGTCTTTC
1401  TCATACTCGA GGGAAATTCG ACACAAACAG CGGAGAAGTG TGGCTAAACC
1451  GGCAAGTGCC TGCAAGATCC ACAGAACTAA CCGCACGAAC TGGCGGTCAG
1501  AAAAGAGCCT GTTCCGGAAA GAGAGAAACA GAGAAACGAT CATGATGGGA
1551  AAGCGGGGAT TCGGCGAAGA ACGAGACTGG AAAGGGAAAA AGAGAAATAC
1601  TGGTGGAAGT ATTCGGACCT TTGGCGAAGT CCGAACCCTT GAAACCCAAA
1651  GATGATCGAT GATTCATTTT TCAATGCGCT ACGGTTCCTG CCGCTCGTGG
1701  GAACCCCACG CAAAACATAT TATTCGCTTC TCTCTGCTGA CAACTCCGGT
1751  TTACGTTATA CCGTATTAGG ATCACTATAA GGGTTCCTTC GGGAGGAGGG
1801  GGGAGGGGAA GAATGTACAT CGTCATAAGG CCTTTATGGT GTGAAGTGGG
1851  TTTTGCGTGG AAAATTCGTT TTCAATGATA TAGAGCCCAC GCATATACGT
1901  ACATACTAGT GGCCAAAAGC GTGGGGTGGG CGGACAAAGC TACACTGGTA
1951  AAATACAGGA TTCTATGAAC AATAACAACA ACCAGCTCAC GTTGCTGAAC
2001  AGCCGAGGTC AGCCGATGCA ACCGAGGTTT CCAAAGTAGC ATTTCTGTGC
2051  TAGCTATGTC TGTAGGTTTA CATTTAATGG TGCGTGGTTC CAGCTTCATG
2101  TGCTTGCATG TGATGTCCTG CAGATGGTAA GAAGATTCTG AAAGCCGCGC
2151  TAGGAGAAAA ATATTCTGCT CGAAGATCTG TCCTCTTAAG TAGAAAGCGT
2201  GAAATTGTTG CGTTCTTGCA TTACTACTCA ACGCGTACGC AAATGCGTCT
2251  ACTGCACCTG CATGATAAAG CTTATGTATC AAAAATTTAA CATCTTGAAA
2301  ATACACAAGT GGTGCAAAGA TGTGTCACGT TCTGGACCTG AGTGGTGCCA
2351  TGTATGCTAT TTAACATGCA AAGGGGAAGA CCCTTCCGCC TTACTGCAAT
2401  AATAAAAAGT ATTTTACGCG TTACCCAATA TAGCAAAGTT TCGCGCAAAA
2451  AAAAAAATAA AAAACAATTA CAAACAAAAA GAAAAAAAAG GAAATAATAG
2501  AAGATCTAAC TGAAGCGAAG GCCAAAACTC TTCTCACTTG ACGTAATAGC
```

```
2551  CGATACAAAA TCTAGAGCAG CAACTTTTCT CTTTCTTCAC TAAAGCTGCT
2601  ACGAAAGTAT AGAAAAATCA AACGCTCAGA ACTTAGCTCT ATTTCAAGGT
2651  ACCATATATA TTTCCTTATA ACTGATGTTA ATTAACTCTA AAGGTGAAGA
2701  ATTATTCACT GGTGTTGTCC CAATTTTGGT TGAATTAGAT GGTGATGTTA
2751  ATGGTCACAA ATTTTCTGTC TCCGGTGAAG GTGAAGGTGA TGCTACTTAC
2801  GGTAAATTGA CCTTAAAATT TATTTGTACT ACTGGTAAAT TGCCAGTTCC
2851  ATGGCCAACC TTAGTCACTA CTTTCGGTTA TGGTGTTCAA TGTTTTGCGA
2901  GATACCCAGA TCATATGAAA CAACATGACT TTTTCAAGTC TGCCATGCCA
2951  GAAGGTTATG TTCAAGAAAG AACTATTTTT TTCAAAGATG ACGGTAACTA
3001  CAAGACCAGA GCTGAAGTCA AGTTTGAAGG TGATACCTTA GTTAATAGAA
3051  TCGAATTAAA AGGTATTGAT TTTAAAGAAG ATGGTAACAT TTTAGGTCAC
3101  AAATTGGAAT ACAACTATAA CTCTCACAAT GTTTACATCA TGGCTGACAA
3151  ACAAAAGAAT GGTATCAAAG TTAACTTCAA AATTAGACAC AACATTGAAG
3201  ATGGTTCTGT TCAATTAGCT GACCATTATC AACAAAATAC TCCAATTGGT
3251  GATGGTCCAG TCTTGTTACC AGACAACCAT TACTTATCCA CTCAATCTGC
3301  CTTATCCAAA GATCCAAACG AAAAGAGAGA CCACATGGTC TTGTTAGAAT
3351  TTGTTACTGC TGCTGGTATT ACCCATGGTA TGGATGAATT GTACAAATAA
3401  CTGCAGGGCG CGCCACTTCT AAATAAGCGA ATTTCTTATG ATTTATGATT
3451  TTTATTATTA AATAAGTTAT AAAAAAAATA AGTGTATACA AATTTTAAAG
3501  TGACTCTTAG GTTTTAAAAC GAAAATTCTT ATTCTTGAGT AACTCTTTCC
3551  TGTAGGTCAG GTTGCTTTCT CAGGTATAGT ATGAGGTCGC TCTTATTGAC
3601  CACACCTCTA CCGGCAGATC CGCTAGGGAT AACAGGGTAA TATAGATCTG
3651  CCCGCCGGGA AGGCGAACCC GATCGGATGC ATCCTCTCTG CTGCCATGAT
3701  GCTGAAGTTG TCGTTGAACA TGGTTGCTGC CGGCGAGGCG GTCGAGCAGG
3751  CAGTGCAGGA GGTGTTGGAC TCGGGAGTCA GAACGGGCGA CCTGCTCGGC
3801  TCGAGCTCGA ATTCATCGAT GATATCAGAT CCACTAGTGG CCTATGCGGC
3851  CGCGGATCTG CCGGTCTCCC TATAGTGAGT CGTATTAATT TCGATAAGCC
3901  AGGTTAACCT GCATTAATGA ATCGGCCAAC GCGCGGGGAG AGGCGGTTTG
3951  CGTATTGGGC GCTCTTCCGC TTCCTCGCTC ACTGACTCGC TGCGCTCGGT
4001  CGTTCGGCTG CGGCGAGCGG TATCAGCTCA CTCAAAGGCG GTAATACGGT
4051  TATCCACAGA ATCAGGGGAT AACGCAGGAA AGAACATGTG AGCAAAAGGC
4101  CAGCAAAAGG CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTTTCCA
4151  TAGGCTCCGC CCCCCTGACG AGCATCACAA AAATCGACGC TCAAGTCAGA
4201  GGTGGCGAAA CCCGACAGGA CTATAAAGAT ACCAGGCGTT TCCCCCTGGA
4251  AGCTCCCTCG TGCGCTCTCC TGTTCCGACC CTGCCGCTTA CCGGATACCT
4301  GTCCGCCTTT CTCCCTTCGG GAAGCGTGGC GCTTTCTCAA TGCTCACGCT
4351  GTAGGTATCT CAGTTCGGTG TAGGTCGTTC GCTCCAAGCT GGGCTGTGTG
4401  CACGAACCCC CCGTTCAGCC CGACCGCTGC GCCTTATCCG GTAACTATCG
4451  TCTTGAGTCC AACCCGGTAA GACACGACTT ATCGCCACTG GCAGCAGCCA
4501  CTGGTAACAG GATTAGCAGA GCGAGGTATG TAGGCGGTGC TACAGAGTTC
4551  TTGAAGTGGT GGCCTAACTA CGGCTACACT AGAAGGACAG TATTTGGTAT
4601  CTGCGCTCTG CTGAAGCCAG TTACCTTCGG AAAAAGAGTT GGTAGCTCTT
4651  GATCCGGCAA ACAAACCACC GCTGGTAGCG GTGGTTTTTT TGTTTGCAAG
4701  CAGCAGATTA CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT
4751  TTCTACGGGG TCTGACGCTC AGTGGAACGA AAACTCACGT TAAGGGATTT
4801  TGGTCATGAG ATTATCAAAA AGGATCTTCA CCTAGATCCT TTTAAATTAA
4851  AAATGAAGTT TTAAATCAAT CTAAAGTATA TATGAGTAAA CTTGGTCTGA
4901  CAGTTACCAA TGCTTAATCA GTGAGGCACC TATCTCAGCG ATCTGTCTAT
4951  TTCGTTCATC CATAGTTGCC TGACTCCCCG TCGTGTAGAT AACTACGATA
5001  CGGGAGGGCT TACCATCTGG CCCCAGTGCT GCAATGATAC CGCGAGACCC
5051  ACGCTCACCG GCTCCAGATT TATCAGCAAT AAACCAGCCA GCCGGAAGGG
5101  CCGAGCGCAG AAGTGGTCCT GCAACTTTAT CCGCCTCCAT CCAGTCTATT
5151  AATTGTTGCC GGGAAGCTAG AGTAAGTAGT TCGCCAGTTA ATAGTTTGCG
```

```
5201  CAACGTTGTT GCCATTGCTA CAGGCATCGT GGTGTCACGC TCGTCGTTTG
5251  GTATGGCTTC ATTCAGCTCC GGTTCCCAAC GATCAAGGCG AGTTACATGA
5301  TCCCCCATGT TGTGCAAAAA AGCGGTTAGC TCCTTCGGTC CTCCGATCGT
5351  TGTCAGAAGT AAGTTGGCCG CAGTGTTATC ACTCATGGTT ATGGCAGCAC
5401  TGCATAATTC TCTTACTGTC ATGCCATCCG TAAGATGCTT TTCTGTGACT
5451  GGTGAGTACT CAACCAAGTC ATTCTGAGAA TAGTGTATGC GGCGACCGAG
5501  TTGCTCTTGC CCGGCGTCAA TACGGGATAA TACCGCGCCA CATAGCAGAA
5551  CTTTAAAAGT GCTCATCATT GGAAAACGTT CTTCGGGGCG AAAACTCTCA
5601  AGGATCTTAC CGCTGTTGAG ATCCAGTTCG ATGTAACCCA CTCGTGCACC
5651  CAACTGATCT TCAGCATCTT TTACTTTCAC CAGCGTTTCT GGGTGAGCAA
5701  AAACAGGAAG GCAAAATGCC GCAAAAAAGG GAATAAGGGC GACACGGAAA
5751  TGTTGAATAC TCATACTCTT CCTTTTTCAA TATTATTGAA GCATTTATCA
5801  GGGTTATTGT CTCATGAGCG GATACATATT TGAATGTATT TAGAAAAATA
5851  AACAAATAGG GGTTCCGCGC ACATTTCCCC GAAAAGTGCC ACCTGACGTC
5901  GAATATCATT GAGAAGCTGC ATTTTTTTTT TTTTTTTTTT TTTTTTTTTT
5951  TATATATATT TCAAGGATAT ACCATTGTAA TGTCTGCCCC TAAGAAGATC
6001  GTCGTTTTGC CAGGTGACCA CGTTGGTCAA GAAATCACAG CCGAAGCCAT
6051  TAAGGTTCTT AAAGCTATTT CTGATGTTCG TTCCAATGTC AAGTTCGATT
6101  TCGAAAATCA TTTAATTGGT GGTGCTGCTA TCGATGCTAC AGGTGTTCCA
6151  CTTCCAGATG AGGCGCTGGA AGCCTCCAAG AAGGCTGATG CCGTTTTGTT
6201  AGGTGCTGTG GGTGGTCCTA AATGGGGTAC CGGTAGTGTT AGACCTGAAC
6251  AAGGTTTACT AAAAATCCGT AAAGAACTTC AATTGTACGC CAACTTAAGA
6301  CCATGTAACT TTGCATCCGA CTCTCTTTTA GACTTATCTC CAATCAAGCC
6351  ACAATTTGCT AAAGGTACTG ACTTCGTTGT TGTCAGAGAA TTAGTGGGAG
6401  GTATTTACTT TGGTAAGAGA AAGGAAGACG ATGGTGATGG TGTCGCTTGG
6451  GATAGTGAAC AATACACCGT TCCAGAAGTG CAAAGAATCA CAAGAATGGC
6501  CGCTTTCATG GCCCTACAAC ATGAGCCACC ATTGCCTATT TGGTCCTTGG
6551  ATAAAGCTAA TGTTTTGGCC TCTTCAAGAT TATGGAGAAA AACTGTGGAG
6601  GAAACCATCA AGAACGAATT CCCTACATTG AAGGTTCAAC ATCAATTGAT
6651  TGATTCTGCC GCCATGATCC TAGTTAAGAA CCCAACCCAC CTAAATGGTA
6701  TTATAATCAC CAGCAACATG TTTGGTGATA TCATCTCCGA TGAAGCCTCC
6751  GTTATCCCAG GTTCCTTGGG TTTGTTGCCA TCTGCGTCCT TGGCCTCTTT
6801  GCCAGACAAG AACACCGCAT TTGGTTTGTA CGAACCATGC CACGGTTCTG
6851  CTCCAGATTT GCCAAAGAAT AAGGTCAACC CTATCGCCAC TATCTTGTCT
6901  GCTGCAATGA TGTTGAAATT GTCATTGAAC TTGCCTGAAG AAGGTAAGGC
6951  CATTGAAGAT GCAGTTAAAA AGGTTTTGGA TGCAGGTATC AGAACTGGTG
7001  ATTTAGGTGG TTCCAACAGT ACCACCGAAG TCGGTGATGC TGTCGCCGAA
7051  GAAGTTAAGA AAATCCTTGC TTAAAAAGAT TCTCTTTTTT TATGATATTT
7101  GTACAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
7151  AAAATGCAGC GTCACATCGG ATAATAATGA CGTCTAAGAA ACCATTATTA
7201  TCATGACATT AACCTATAAA AATAGGCGTA TCACGAGGCC CTTTCGTCTC
7251  GCGCGTTTCG GTGATGACGG TGAAAACCTC TGACACATGC AGCTCCCGGA
7301  GACGGTCACA GCTTGTCTGT AAGCGGATGC CGGGAGCAGA CAAGCCCGTC
7351  AGGGCGCGTC AGCGGGTGTT GGCGGGTGTC GGGGCTGGCT TAACTATGCG
7401  GCATCAGAGC AGATTGTACT GAGAGTGCAC CATATGGACA TATTGTCGTT
7451  AGAACGCGGC TACAATTAAT ACATAACCTT ATGTATCATA CACATACGAT
7501  TTAGGTGACA CTATA
//
```

Fig:42 pGenIn022A - 12093 bp

```
   1  GAACGCGGCC GCCAGCTGAA GCTTCGTACG CTGCAGGTCG ACGGATCAAA
  51  ATTGTGACAG CTTTCCAGAA TGGATTATTT TTCCTCAAAT TCCTTGTCTT
 101  CCTGTTTTCA TCTGGACCAT CTCCATAATG AAGCCTTACA TGTTTGGCAC
 151  GTAGCGGAAC GTGATCGTCA CAAACCGTAA GGTAGAGACC CCAGATTTTC
 201  GCATTTTCTC TTAAACTCTC CATTAGCTTA GGATCTGACG ATCACCTAGC
 251  GACTCTCTCC ACCGTTTGAC GAGGCCATTT ACAAAAACAT AACGAACGAC
 301  AAGCCTACTC GAATTCGTTT CCAAACTCTT TTCGAACTTG TCTTCAACTG
 351  CTTTCGCATG AAGTACCTCC CAACTACTTT TCCTCACACT TGTACTCCAT
 401  GACTAAACCC CCCCTCCCAT TACAAACTAA AATCTTACTT TTATTTTCTT
 451  TTGCCCTCTC TGTCGCTCTG CCTTAACTAC GTATTTCTCG CCGAGAAAAA
 501  CTTCAATTTA AGCTATTCTC CAAAAATCTT AGCGTATATT TTTTTTCCAA
 551  AGTGACAGGT GCCCCGGGTA ACCCAGTTCC TCACTATTTT TTACTGCGGA
 601  AGCGGAAGCG GAAAATACGG AAACGCGCGG GAACATACAA AACATACAAA
 651  ATATACCTTT CTCACACAAG AAATATATGC TACTTGCAAA ATATCATACC
 701  AAAAAACTTT TCACAACCGA AACCAAAACC AACGGATATC ATACATTACA
 751  CTACCACCAT TCAAACTTTA CTACTATCCT CCCTTCAGTT TCCCTTTTTC
 801  TGCCTTTTTC GGTGACGGAA ATACGCTTCA GAGACCCTAA AGGGAAATCC
 851  ATGCCATAAC AGGAAAGTAA CATCCCAATG CGGACTATAC CACCCCACCA
 901  CACTCCTACC AATAACGGTA ACTATTCTAT GTTTTCTTAC TCCTATGTCT
 951  ATTCATCTTT CATCTGACTA CCTAATACTA TGCAAAAATG TAAAATCATC
1001  ACACAAAACA TAAACAATCA AAATCAGCCA TTTCCGCACC TTTTCCTCTG
1051  TCCACTTTCA ACCGTCCCTC CAAATGTAAA ATGGCCTATC GGAATACATT
1101  TTCTACATCC TAACTACTAT AAAACAACCT TTAGACTTAC GTTTGCTACT
1151  CTCATGGTCT CAATACTGCC GCCGACATTC TGTCCCACAT ACTAAATCTC
1201  TTCCCGTCAT TATCGCCCGC ATCCGGTGCC GTAAATGCAA AACAAATACC
1251  ATCTATGTCT TCCACACCAT CATTTTACTA TGCCTGCCAC CATCCATTTG
1301  TCTTTTGCAC CATATCTTCA TAACCTGTCA CCTTGAAACT ACCTCTGCAT
1351  GCCACCTACC GACCAACTTT CATGTTCTGT TTCGACCTAC CTCTTGTAAA
1401  TGACAAATCA CCTTTTTCAT CGTATGCACC TTATTCTCCA CATCACAATG
1451  CACTATTGCT TTTGCTTTTT CACCTGTCAT ATCCTATTGC TATTAGATGA
1501  AATATAATAA AAATTGTCCT CCACCCATAA CACCTCTCAC TCCCACCTAC
1551  TGAACATGTC TGGACCCTGC CCTCATATCA CCTGCGTTTC CGTTAAACTA
1601  TCGGTTGCGG CCATATCTAC CAGAAAGCAC CGTTTCCCGT CCGATCAACT
1651  GTAGTTAAGC TGGTAAGAGC CTGACCGAGT AGTGTAGTGG GTGACCATAC
1701  GCGAAACTCA GGTGCTGCAA TCTTTATTTC TTTTTTTTTT TTTTTTTTTT
1751  TTTTTTTTTC TAGTTTCTTG GCTTCCTATG CTAAATCCCA TAACTAACCT
1801  ACCATTCGAT TCAGAAAAAT TCGCACTATC CAGCTGCACT CTTCTTCTGA
1851  AGAGTTAAGC ACTCCATTAT GCTCATTGGG TTGCTACTAC TTGATATGTA
1901  CAAACAATAT TCTCCTCCGA TATTCCTACA AAAAAAAAAA AAAAAACACT
1951  CCGGTTTTGT TCTCTTCCCT CCATTTCCCT CTCTTCTACG GTTAATACTT
2001  TCCTCTTCGT CTTTTTCTAC ACCCTCGTTT AGTTGCTTCT TATTCCTTCC
2051  CGCTTTCCTG CACTAACATT TTGCCGCATT ACACTATATG ATCGTAGTAC
2101  ATCTTACAAC TCCGCATACC GCGTCGCCGC GTCGCCGCGT CGCCAAAAAT
2151  TTACTTCGCC AACCATTCCA TATCTGTTAA GTATACATGT ATATATTGCA
2201  CTGGCTATTC ATCTTGCACT TTTCCTCTTT CTTCTTCCCA GTAGCCTCAT
2251  CCTTTTACGC TGCCTCTCTG GAACTTGCCA TCATCATTCC CTAGAAACTG
2301  CCATTTACTT AAAAAAAAAA AAAAAAAAAA AATGTCCCCA CTGTTCACTG
2351  TTCACTGTTC ACTTGTCTCT TACATCTTTC TTGGTAAAAT CGTAGTTCGT
2401  AGTATTTTTT TTCATATCAA AGGCATGTCC TGTTAACTAT AGGAAATGAG
2451  CTTTTCTCAA TTCTCTAAAC TTATACAAGC ACCTCATGTT TGCCGCTCTG
```

```
2501  ATGGTGCGGA AAAAACTGCT CCATGAAGCA AACTGTCCGG GCAAATCCTT
2551  TCACGCTCGG GAAGCTTTGT GAAAGCCCTT CTCTTTCAAC CCATCTTTGC
2601  AACGAAAAAA AAAAAAAAAA TAAAAAATAA AAAGACCAAA TAGTAAATAG
2651  TAACTTACAT ACATTAGTAA ATGGTACACT CTTACACACT ATCATCCTCA
2701  TCGTATATTA TAATAGATAT ATACAATACA TGTTTTTACC CGGATCATAG
2751  AATTCTTAAG ACAAATAAAA TTTATAGAGA CTTGTTCAGT CTACTTCTCT
2801  CTAAACTAGG CCCCGGCTCC TGCCAGTACC CACTTAGAAA GAAATAAAAA
2851  ACAAATCAGA CAACAAAGGC TTAATCTCAG CAGATCGTAA CAACAAGGCT
2901  ACTCTACTGC TTACAATACC CCGTTGTACA TCTAAGTCGT ATACAAATGA
2951  TTTATCCCCA CGCAAAATGA CATTGCAATT CGCCAGCAAG CACCCAAGGC
3001  CTTTCCGCCA AGTGCACCGT TGCTAGCCTG CTATGGTTCA GCGACGCCAC
3051  AAGGACGCCT TATTCGTATC CATCTATATT GTGTGGAGCA AAGAAATCAC
3101  CGCGTTCTAG CATGGATTCT GACTTAGAGG CGTTCAGCCA TAATCCAGCG
3151  GATGGTAGCT TCGCGGCAAT GCCTGATCAG ACAGCCGCAA AAACCAATTA
3201  TCCGAATGAA CTGTTCCTCT CGTACTAAGT TCAATTACTA TTGCGGTAAC
3251  ATTCATCAGT AGGGTAAAAC TAACCTGTCT CACGACGGTC TAAACCCAGC
3301  TCACGTTCCC TATTAGTGGG TGAACAATCC AACGCTTACC GAATTCTGCT
3351  TCGGTATGAT AGGAAGAGCC GACATCGAAG AATCAAAAAG CAATGTCGCT
3401  ATGAACGCTT GACTGCCACA AGCCAGTTAT CCCTGTGGTA ACTTTTCTGG
3451  CACCTCTAGC CTCAAATTCC GAGGGACTAA AGGATCGATA GGCCACACTT
3501  TCATGGTTTG TATTCACACT GAAAATCAAA ATCAAGGGGG CTTTTACCCT
3551  TTTGTTCTAC TGGAGATTTC TGTTCTCCAT GAGCCCCCCT TAGGACATCT
3601  GCGTTATCGT TTAACAGATG TGCCGCCCCA GCCAAACTCC CCACCTGACA
3651  ATGTCTTCAA CCCGGATCAG CCCCGAATGG GACCTTGAAT GCTAGAACGT
3701  GGAAAATGAA TTCCAGCTCC GCTTCATTGA ATAAGTAAAG AAACTATAAA
3751  GGTAGTGGTA TTTCACTGGC GCCGAAGCTC CCACTTATTC TACACCCTCT
3801  ATGTCTCTTC ACAATGTCAA ACTAGAGTCA AGCTCAACAG GGTCTTCTTT
3851  CCCCGCTGAT TCTGCCAAGC CCGTTCCCTT GGCTGTGGTT TCGCTAGATA
3901  GTAGATAGGG ACAGTGGGAA TCTCGTTAAT CCATTCATGC GCGTCACTAA
3951  TTAGATGACG AGGCATTTGG CTACCTTAAG AGAGTCATAG TTACTCCCGC
4001  CGTTTACCCG CGCTTGGTTG AATTTCTTCA CTTTGACATT CAGAGCACTG
4051  GGCAGAAATC ACATTGCGTC AACATCACTT TCTGACCATC GCAATGCTAT
4101  GTTTTAATTA GACAGTCAGA TTCCCCTTGT CCGTACCAGT TCTAAGTTGA
4151  TCGTTAATTG TAGCAAGCGA CGGTCTACAA GAGACCTACC AAGGCCGTCT
4201  ACAACAAGGC ACGCAAGTAG TCCGCCTAGC AGAGCAAGCC CCACCAAGCA
4251  GTCCACAAGC ACGCCCGCTG CGTCTGACCA AGGCCCTCAC TACCCGACCC
4301  TTAGAGCCAA TCCTTATCCC GAAGTTACGG ATCTATTTTG CCGACTTCCC
4351  TTATCTACAT TATTCTATCA ACTAGAGGCT GTTCACCTTG GAGACCTGCT
4401  GCGGTTATCA GTACGACCTG GCATGAAAAC TATTCCTTCC TGTGGATTTT
4451  CACGGGCCGT CACAAGCGCA CCGGAGCCAG CAAAGGTGCT GGCCTCTTCC
4501  AGCCATAAGA CCCCATCTCC GGATAAACCA ATTCCGGGGT GATAAGCTGT
4551  TAAGAAGAAA AGATAACTCC TCCCAGGGCT CGCGCCGACG TCTCCACATT
4601  CAGTTACGTT ACCGTGAAGA ATCCATATCC AGGTTCCGGA ATCTTAACCG
4651  GATTCCCTTT CGATGGTGGC CTGCATAAAA TCAGGCCTTT GAAACGGAGC
4701  TTCCCCATCT CTTAGGATCG ACTAACCCAC GTCCAACTGC TGTTGACGTG
4751  GAACCTTTCC CCACTTCAGT CTTCAAAGTT CTCATTTGAA TATTTGCTAC
4801  TACCACCAAG ATCCAAGCTA TCTACTGAGA TTTCTGGCTC TTTTGTTGTA
4851  CTGTCACCTA ACCACAGACC AAGCATCCAA GCCATACTTT TTACAGCAGG
4901  AGTTACAAGG TCACTACGTC CAGTGAGAAA TTTAGATAAA ACACCATTTC
4951  CTGCGAGTAC TGGACCAAAT CTTATGCAGC TAGAAATTCT CAATTGAGCA
5001  TCAAGATAAT CCAAATCTCT AACTTCAATG TCAAAGTTGA AATATTCTCC
5051  TTTAGAGCGC TCCATTTCTT CTATGAAGCG TTTTGCGGCA AACTCACCTT
5101  CAACTGTCAT TGGGAATGTC TTATGATGGT TTTTTGGAAT TATTATTATC
```

```
5151  CTACCATCAA GCGTCTGACA TTGCTGCAGA TTTCTCCATC TCACTTTATA
5201  TTTGGTGGCA TTTCTACCAC TTTTTTCCAA CAGTGGTTTG GTAGGGACCC
5251  TGACTGACAA TTTATGACCT GCAGTACATT GTAATGCAAG ACGCTGATAA
5301  ACTGTTCTAC GCCTGGGATC TAACCTACCA GGTTCACCTT CAAAAGCTCT
5351  GTGTTTGGTT TTTTGCTGTA TATTATAGAT TTTCTGATAG CCCTGTGTGA
5401  CATTTATGAC GCGGGCAGCG GAGCCATCTG CGCACATAAC GTAAGAGTTA
5451  GCCGTGACGT TTGCGATGTC TTTAATTTCA CCGTTAGCCA TCAGAATAGT
5501  CGTGTTTTCA GAAAGCATTT TGATCCGACA TACGATGACC TCAATGATTT
5551  AGATTATGTG TTGCACTTTT ATAGACCTAC CAAAAATCCA GTGCGTACAC
5601  TAATACTTTC ATAAAGATAC CTGAAACAAT AACCAGAAAG ATCGGCAAAA
5651  AAATTTTTTT TCTTTGCCGA GATCACAAAC CTACTATGAC GAAAAAGCTT
5701  GAAGTTTAGA TGAGTAAGGA AAATACAAGT GACGCTTTTA TATGGTGCAA
5751  GGAACAAAAA CTAAAAACAA CAAGGCAAAT GTGGATCTGT CATGTATGGC
5801  AACGACAGCA GGATGGCTCA CAAAAAAAGA CAAAAAAAAC TAAGGCAAAA
5851  GAACAAAGCT CCTCTCCTGC TCAAGAAACG TATTGTTGAA AAACCACCGT
5901  CGTAAGAAAG TTTTTCTGTG ACCTATAATG GTTTAAAATC GGCCCATTTT
5951  TTTTCCCTCT TTTGTGGTCC AGTCTTTCTC ATACTCGAGG GAAATTCGAC
6001  ACAAACAGCG GAGAAGTGTG GCTAAACCGG CAAGTGCCTG CAAGATCCAC
6051  AGAACTAACC GCACGAACTG GCGGTCAGAA AAGAGCCTGT TCCGGAAAGA
6101  GAGAAACAGA GAAACGATCA TGATGGGAAA GCGGGGATTC GGCGAAGAAC
6151  GAGACTGGAA AGGGAAAAAG AGAAATACTG GTGGAAGTAT TCGGACCTTT
6201  GGCGAAGTCC GAACCCTTGA AACCCAAAGA TGATCGATGA TTCATTTTTC
6251  AATGCGCTAC GGTTCCTGCC GCTCGTGGGA ACCCCACGCA AAACATATTA
6301  TTCGCTTCTC TCTGCTGACA ACTCCGGTTT ACGTTATACC GTATTAGGAT
6351  CACTATAAGG GTTCCTTCGG GAGGAGGGGG GAGGGGAAGA ATGTACATCG
6401  TCATAAGGCC TTTATGGTGT GAAGTGGGTT TTGCGTGGAA AATTCGTTTT
6451  CAATGATATA GAGCCCACGC ATATACGTAC ATACTAGTGG CCAAAAGCGT
6501  GGGGTGGGCG GACAAAGCTA CACTGGTAAA ATACAGGATT CTATGAACAA
6551  TAACAACAAC CAGCTCACGT TGCTGAACAG CCGAGGTCAG CCGATGCAAC
6601  CGAGGTTTCC AAAGTAGCAT TTCTGTGCTA GCTATGTCTG TAGGTTTACA
6651  TTTAATGGTG CGTGGTTCCA GCTTCATGTG CTTGCATGTG ATGTCCTGCA
6701  GATGGTAAGA AGATTCTGAA AGCCGCGCTA GGAGAAAAAT ATTCTGCTCG
6751  AAGATCTGTC CTCTTAAGTA GAAAGCGTGA AATTGTTGCG TTCTTGCATT
6801  ACTACTCAAC GCGTACGCAA ATGCGTCTAC TGCACCTGCA TGATAAAGCT
6851  TATGTATCAA AAATTTAACA TCTTGAAAAT ACACAAGTGG TGCAAAGATG
6901  TGTCACGTTC TGGACCTGAG TGGTGCCATG TATGCTATTT AACATGCAAA
6951  GGGGAAGACC CTTCCGCCTT ACTGCAATAA TAAAAAGTAT TTTACGCGTT
7001  ACCCAATATA GCAAAGTTTC GCGCAAAAAA AAAAATAAAA AACAATTACA
7051  AACAAAAAGA AAAAAAAGGA AATAATAGAA GATCTAACTG AAGCGAAGGC
7101  CAAAACTCTT CTCACTTGAC GTAATAGCCG ATACAAAATC TAGAGCAGCA
7151  ACTTTTCTCT TTCTTCACTA AAGCTGCTAC GAAAGTATAG AAAAATCAAA
7201  CGCTCAGAAC TTAGCTCTAT TTCAAGGTAC CATATATATT TCCTTATAAC
7251  TGATGTTAAT TAACTCTAAA GGTGAAGAAT TATTCACTGG TGTTGTCCCA
7301  ATTTTGGTTG AATTAGATGG TGATGTTAAT GGTCACAAAT TTTCTGTCTC
7351  CGGTGAAGGT GAAGGTGATG CTACTTACGG TAAATTGACC TTAAAATTTA
7401  TTTGTACTAC TGGTAAATTG CCAGTTCCAT GGCCAACCTT AGTCACTACT
7451  TTCGGTTATG GTGTTCAATG TTTTGCGAGA TACCCAGATC ATATGAAACA
7501  ACATGACTTT TTCAAGTCTG CCATGCCAGA AGGTTATGTT CAAGAAAGAA
7551  CTATTTTTTT CAAAGATGAC GGTAACTACA AGACCAGAGC TGAAGTCAAG
7601  TTTGAAGGTG ATACCTTAGT TAATAGAATC GAATTAAAAG GTATTGATTT
7651  TAAAGAAGAT GGTAACATTT TAGGTCACAA ATTGGAATAC AACTATAACT
7701  CTCACAATGT TTACATCATG GCTGACAAAC AAAAGAATGG TATCAAAGTT
7751  AACTTCAAAA TTAGACACAA CATTGAAGAT GGTTCTGTTC AATTAGCTGA
7801  CCATTATCAA CAAAATACTC CAATTGGTGA TGGTCCAGTC TTGTTACCAG
```

```
 7851  ACAACCATTA CTTATCCACT CAATCTGCCT TATCCAAAGA TCCAAACGAA
 7901  AAGAGAGACC ACATGGTCTT GTTAGAATTT GTTACTGCTG CTGGTATTAC
 7951  CCATGGTATG GATGAATTGT ACAAATAACT GCAGGGCGCG CCACTTCTAA
 8001  ATAAGCGAAT TTCTTATGAT TTATGATTTT TATTATTAAA TAAGTTATAA
 8051  AAAAAATAAG TGTATACAAA TTTTAAAGTG ACTCTTAGGT TTTAAAACGA
 8101  AAATTCTTAT TCTTGAGTAA CTCTTTCCTG TAGGTCAGGT TGCTTTCTCA
 8151  GGTATAGTAT GAGGTCGCTC TTATTGACCA CACCTCTACC GGCAGATCCG
 8201  CTAGGGATAA CAGGGTAATA TAGATCTGCC CGCCGGGAAG GCGAACCCGA
 8251  TCGGATGCAT CCTCTCTGCT GCCATGATGC TGAAGTTGTC GTTGAACATG
 8301  GTTGCTGCCG GCGAGGCGGT CGAGCAGGCA GTGCAGGAGG TGTTGGACTC
 8351  GGGAGTCAGA ACGGGCGACC TGCTCGGCTC GAGCTCGAAT TCATCGATGA
 8401  TATCAGATCC ACTAGTGGCC TATGCGGCCG CGGATCTGCC GGTCTCCCTA
 8451  TAGTGAGTCG TATTAATTTC GATAAGCCAG GTTAACCTGC ATTAATGAAT
 8501  CGGCCAACGC GCGGGGAGAG GCGGTTTGCG TATTGGGCGC TCTTCCGCTT
 8551  CCTCGCTCAC TGACTCGCTG CGCTCGGTCG TTCGGCTGCG GCGAGCGGTA
 8601  TCAGCTCACT CAAAGGCGGT AATACGGTTA TCCACAGAAT CAGGGGATAA
 8651  CGCAGGAAAG AACATGTGAG CAAAAGGCCA GCAAAAGGCC AGGAACCGTA
 8701  AAAAGGCCGC GTTGCTGGCG TTTTTCCATA GGCTCCGCCC CCCTGACGAG
 8751  CATCACAAAA ATCGACGCTC AAGTCAGAGG TGGCGAAACC CGACAGGACT
 8801  ATAAAGATAC CAGGCGTTTC CCCCTGGAAG CTCCCTCGTG CGCTCTCCTG
 8851  TTCCGACCCT GCCGCTTACC GGATACCTGT CCGCCTTTCT CCCTTCGGGA
 8901  AGCGTGGCGC TTTCTCAATG CTCACGCTGT AGGTATCTCA GTTCGGTGTA
 8951  GGTCGTTCGC TCCAAGCTGG GCTGTGTGCA CGAACCCCCC GTTCAGCCCG
 9001  ACCGCTGCGC CTTATCCGGT AACTATCGTC TTGAGTCCAA CCCGGTAAGA
 9051  CACGACTTAT CGCCACTGGC AGCAGCCACT GGTAACAGGA TTAGCAGAGC
 9101  GAGGTATGTA GGCGGTGCTA CAGAGTTCTT GAAGTGGTGG CCTAACTACG
 9151  GCTACACTAG AAGGACAGTA TTTGGTATCT GCGCTCTGCT GAAGCCAGTT
 9201  ACCTTCGGAA AAAGAGTTGG TAGCTCTTGA TCCGGCAAAC AAACCACCGC
 9251  TGGTAGCGGT GGTTTTTTTG TTTGCAAGCA GCAGATTACG CGCAGAAAAA
 9301  AAGGATCTCA AGAAGATCCT TTGATCTTTT CTACGGGGTC TGACGCTCAG
 9351  TGGAACGAAA ACTCACGTTA AGGGATTTTG GTCATGAGAT TATCAAAAAG
 9401  GATCTTCACC TAGATCCTTT TAAATTAAAA ATGAAGTTTT AAATCAATCT
 9451  AAAGTATATA TGAGTAAACT TGGTCTGACA GTTACCAATG CTTAATCAGT
 9501  GAGGCACCTA TCTCAGCGAT CTGTCTATTT CGTTCATCCA TAGTTGCCTG
 9551  ACTCCCCGTC GTGTAGATAA CTACGATACG GGAGGGCTTA CCATCTGGCC
 9601  CCAGTGCTGC AATGATACCG CGAGACCCAC GCTCACCGGC TCCAGATTTA
 9651  TCAGCAATAA ACCAGCCAGC CGGAAGGGCC GAGCGCAGAA GTGGTCCTGC
 9701  AACTTTATCC GCCTCCATCC AGTCTATTAA TTGTTGCCGG GAAGCTAGAG
 9751  TAAGTAGTTC GCCAGTTAAT AGTTTGCGCA ACGTTGTTGC CATTGCTACA
 9801  GGCATCGTGG TGTCACGCTC GTCGTTTGGT ATGGCTTCAT TCAGCTCCGG
 9851  TTCCCAACGA TCAAGGCGAG TTACATGATC CCCCATGTTG TGCAAAAAAG
 9901  CGGTTAGCTC CTTCGGTCCT CCGATCGTTG TCAGAAGTAA GTTGGCCGCA
 9951  GTGTTATCAC TCATGGTTAT GGCAGCACTG CATAATTCTC TTACTGTCAT
10001  GCCATCCGTA AGATGCTTTT CTGTGACTGG TGAGTACTCA ACCAAGTCAT
10051  TCTGAGAATA GTGTATGCGG CGACCGAGTT GCTCTTGCCC GGCGTCAATA
10101  CGGGATAATA CCGCGCCACA TAGCAGAACT TTAAAAGTGC TCATCATTGG
10151  AAAACGTTCT TCGGGGCGAA AACTCTCAAG GATCTTACCG CTGTTGAGAT
10201  CCAGTTCGAT GTAACCCACT CGTGCACCCA ACTGATCTTC AGCATCTTTT
10251  ACTTTCACCA GCGTTTCTGG GTGAGCAAAA ACAGGAAGGC AAAATGCCGC
10301  AAAAAAGGGA ATAAGGGCGA CACGGAAATG TTGAATACTC ATACTCTTCC
10351  TTTTTCAATA TTATTGAAGC ATTTATCAGG GTTATTGTCT CATGAGCGGA
10401  TACATATTTG AATGTATTTA GAAAAATAAA CAAATAGGGG TTCCGCGCAC
```

```
10451  ATTTCCCCGA AAAGTGCCAC CTGACGTCGA ATATCATTGA GAAGCTGCAT
10501  TTTTTTTTTT TTTTTTTTTT TTTTTTTTTA TATATATTTC AAGGATATAC
10551  CATTGTAATG TCTGCCCCTA AGAAGATCGT CGTTTTGCCA GGTGACCACG
10601  TTGGTCAAGA AATCACAGCC GAAGCCATTA AGGTTCTTAA AGCTATTTCT
10651  GATGTTCGTT CCAATGTCAA GTTCGATTTC GAAAATCATT TAATTGGTGG
10701  TGCTGCTATC GATGCTACAG GTGTTCCACT TCCAGATGAG GCGCTGGAAG
10751  CCTCCAAGAA GGCTGATGCC GTTTTGTTAG GTGCTGTGGG TGGTCCTAAA
10801  TGGGGTACCG GTAGTGTTAG ACCTGAACAA GGTTTACTAA AAATCCGTAA
10851  AGAACTTCAA TTGTACGCCA ACTTAAGACC ATGTAACTTT GCATCCGACT
10901  CTCTTTTAGA CTTATCTCCA ATCAAGCCAC AATTTGCTAA AGGTACTGAC
10951  TTCGTTGTTG TCAGAGAATT AGTGGGAGGT ATTTACTTTG GTAAGAGAAA
11001  GGAAGACGAT GGTGATGGTG TCGCTTGGGA TAGTGAACAA TACACCGTTC
11051  CAGAAGTGCA AAGAATCACA AGAATGGCCG CTTTCATGGC CCTACAACAT
11101  GAGCCACCAT TGCCTATTTG GTCCTTGGAT AAAGCTAATG TTTTGGCCTC
11151  TTCAAGATTA TGGAGAAAAA CTGTGGAGGA AACCATCAAG AACGAATTCC
11201  CTACATTGAA GGTTCAACAT CAATTGATTG ATTCTGCCGC CATGATCCTA
11251  GTTAAGAACC CAACCCACCT AAATGGTATT ATAATCACCA GCAACATGTT
11301  TGGTGATATC ATCTCCGATG AAGCCTCCGT TATCCCAGGT TCCTTGGGTT
11351  TGTTGCCATC TGCGTCCTTG GCCTCTTTGC CAGACAAGAA CACCGCATTT
11401  GGTTTGTACG AACCATGCCA CGGTTCTGCT CCAGATTTGC CAAAGAATAA
11451  GGTCAACCCT ATCGCCACTA TCTTGTCTGC TGCAATGATG TTGAAATTGT
11501  CATTGAACTT GCCTGAAGAA GGTAAGGCCA TTGAAGATGC AGTTAAAAAG
11551  GTTTTGGATG CAGGTATCAG AACTGGTGAT TTAGGTGGTT CCAACAGTAC
11601  CACCGAAGTC GGTGATGCTG TCGCCGAAGA AGTTAAGAAA ATCCTTGCTT
11651  AAAAAGATTC TCTTTTTTTA TGATATTTGT ACAAAAAAAA AAAAAAAAAA
11701  AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AATGCAGCGT CACATCGGAT
11751  AATAATGACG TCTAAGAAAC CATTATTATC ATGACATTAA CCTATAAAAA
11801  TAGGCGTATC ACGAGGCCCT TTCGTCTCGC GCGTTTCGGT GATGACGGTG
11851  AAAACCTCTG ACACATGCAG CTCCCGGAGA CGGTCACAGC TTGTCTGTAA
11901  GCGGATGCCG GGAGCAGACA AGCCCGTCAG GGCGCGTCAG CGGGTGTTGG
11951  CGGGTGTCGG GGCTGGCTTA ACTATGCGGC ATCAGAGCAG ATTGTACTGA
12001  GAGTGCACCA TATGGACATA TTGTCGTTAG AACGCGGCTA CAATTAATAC
12051  ATAACCTTAT GTATCATACA CATACGATTT AGGTGACACT ATA
```